

# Biological Nitrogen Fixation – Simulation of the Reaction Mechanism of Nitrogenase from First Principles

Dissertation  
zur Erlangung des Grades eines Doktors  
der Naturwissenschaften

vorgelegt von  
Johannes Kästner geb. Schimpl  
aus Wien

genehmigt von der Mathematisch-Naturwissenschaftlichen Fakultät  
der Technischen Universität Clausthal

Tag der mündlichen Prüfung  
3. Juni 2004

Die Arbeit wurde angefertigt am  
Institut für Theoretische Physik der Technischen Universität Clausthal

Vorsitzender der Promotionskommission: Prof. Dr. Jürgen Fertig

Hauptberichterstatter: Prof. Dr. Peter E. Blöchl

Berichterstatter: Prof. Dr. Dieter Mayer

## Abstract

Nitrogen is one of the major constituents of biological systems. The shortage of available nitrogen is the reason why fertilizer is used in agriculture. The only relevant inorganic source of nitrogen is our atmosphere. Accessing atmospheric dinitrogen, however, requires to break one of the strongest chemical bonds. For this purpose, nature developed an enzyme of exciting complexity, namely nitrogenase. Despite decades of research, the mechanism of this enzyme remained in the dark.

In this thesis I report on first-principles calculations that aim to unravel the complete catalytic cycle of nitrogen fixation at the FeMo-cofactor of nitrogenase. The calculations describe the electrons, responsible for chemical bonding, quantum mechanically, on the basis of density functional theory and the projector augmented wave method.

The active center of nitrogenase, the so-called FeMo-cofactor, exhibits a complicated spin structure, which has hampered previous calculations. In this work I employ a rather recent development for the description of the spin structure, allowing for non-collinear alignment of the local spins.

By comparing my calculations with available experimental information, I identified the charge state of the cofactor core in its resting state as  $[\text{MoFe}_7\text{S}_9\text{N}]^{\pm 0}$ . My calculations indicate that electrons and protons are transferred in an alternating manner to the cofactor. An investigation of potential proton transport channels through the protein confirmed the existence of a single well-defined proton path.

This work made an attempt of a comprehensive investigation of all possible binding sites for dinitrogen at the cofactor. Binding at the Mo-site has been found not to lead to a viable reaction pathway. Instead I found an axial and a bridged binding mode of dinitrogen to Fe-sites to be relevant. One of the most surprising results of this study was that the cage opens and one protonated sulfur bridge is broken upon binding of dinitrogen. In the un-catalyzed gas phase reaction, the first protonation is the critical time limiting step. In the enzyme,  $\text{N}_2$  is activated by forming already two weak bonds to two Fe-sites. A common theme of the reaction is that the Fe-sites maintain their tetrahedral environment. This is made possible by the presence of a central nitrogen ligand, adds flexibility to the cluster by providing a varying number of bonds to its iron neighbors. Another important observation is that the opening of a sulfur bridge stabilizes the binding of  $\text{N}_2$  and that its closing facilitates desorption of the product.

My findings are supported by my investigation of the reaction of acetylene with the FeMo-cofactor, for which detailed experimental results are available. The mechanism I propose for this reaction explains the experimentally observed stereoselectivity and the role of acetylene as inhibitor of nitrogen fixation. This work lends additional credibility to my proposed reaction mechanism.



# Contents

<b>Abstract</b>	<b>III</b>
<b>1 Introduction</b>	<b>1</b>
<b>I Nitrogenase – a review</b>	<b>5</b>
<b>2 The protein – structure</b>	<b>7</b>
2.1 Proteins of nitrogenase . . . . .	7
2.1.1 The Fe-protein . . . . .	7
2.1.2 The MoFe-protein . . . . .	8
2.1.3 Complex formation . . . . .	10
2.2 Cofactors: electronic properties . . . . .	11
2.2.1 FeMo-cofactor . . . . .	11
2.2.2 P-cluster . . . . .	12
2.2.3 $[\text{Fe}_4\text{S}_4]$ cluster of the Fe-protein . . . . .	13
2.3 Theoretical work concerning the resting state . . . . .	13
2.3.1 Calculations considering the central ligand . . . . .	13
2.3.2 Calculations without considering the central ligand . . . . .	14
<b>3 The mechanism</b>	<b>17</b>
3.1 The Fe-protein cycle . . . . .	17
3.2 The Thorneley-Lowe scheme for nitrogen fixation . . . . .	18
3.2.1 Nitrogen and hydrogen . . . . .	20
3.2.2 Other substrates and inhibitors . . . . .	21
3.3 Mutagenesis experiments . . . . .	24
3.4 Theoretical work concerning the mechanism . . . . .	25
3.4.1 Calculations considering the central ligand . . . . .	26
3.4.2 Calculations without consideration of the central ligand . . . . .	27
<b>4 Open questions</b>	<b>29</b>

<b>II</b>	<b>Theory</b>	<b>31</b>
<b>5</b>	<b>Electronic structure calculations</b>	<b>33</b>
5.1	Density functional theory . . . . .	34
5.1.1	Hohenberg-Kohn theorems . . . . .	34
5.1.2	Kohn-Sham method . . . . .	35
5.1.3	Local density approximation (LDA) . . . . .	37
5.1.4	PBE gradient approximation . . . . .	38
5.2	The projector augmented wave method . . . . .	40
5.2.1	PAW – The overview . . . . .	40
5.2.2	PAW – The mathematical basis . . . . .	41
5.3	Molecular dynamics . . . . .	52
5.3.1	Discretization of the equations of motion . . . . .	52
5.3.2	Ab-initio molecular dynamics . . . . .	53
5.4	Reaction rates . . . . .	54
5.5	Spin . . . . .	55
5.5.1	Quantum mechanics . . . . .	56
5.5.2	Pauli equation . . . . .	56
5.5.3	Spin density functional theory . . . . .	57
5.5.4	Spin contamination . . . . .	58
5.5.5	Collinear vs. non-collinear calculations . . . . .	60
<b>6</b>	<b>Discrete spins – the Heisenberg model</b>	<b>63</b>
6.1	Heisenberg spin Hamiltonian . . . . .	63
6.1.1	Basis set . . . . .	64
6.2	The ground state of a spin-coupled system . . . . .	65
6.2.1	Diagonalization of the Hamilton-matrix . . . . .	65
6.2.2	Iterative search for the ground state . . . . .	66
6.2.3	Analysis . . . . .	68
6.3	Classical spins – the classical Heisenberg model . . . . .	69
6.4	Application: three interacting spins in $[\text{Fe}_3\text{S}_4]^+$ . . . . .	73
6.4.1	The Heisenberg model . . . . .	74
6.4.2	Non-collinear magnetism . . . . .	77
<b>III</b>	<b>Application</b>	<b>79</b>
<b>7</b>	<b>Considerations concerning the structure of nitrogenase</b>	<b>81</b>
7.1	Proton channels . . . . .	81
7.1.1	The proton shuttle mechanism . . . . .	81
7.1.2	Systematic search for proton channels – calculational details . . . . .	84

---

7.1.3	Results . . . . .	85
7.2	Distance from the cofactor to the surface . . . . .	88
7.3	Protonation state of residues . . . . .	90
7.3.1	Protein chains from the P-cluster to the vicinity of the FeMoco 92	
<b>8</b>	<b>Dinitrogen and its reduction products in the gas phase</b>	<b>95</b>
8.1	Electronic structure of dinitrogen . . . . .	95
8.2	Comparison to experiment . . . . .	96
<b>9</b>	<b>The resting state</b>	<b>99</b>
9.1	Determination of the charge state . . . . .	99
9.1.1	Calculational details . . . . .	100
9.1.2	Total spin as criterion for finding the resting state . . . . .	101
9.1.3	Atomic structure as criterion for finding the resting state . . . . .	102
9.2	Analysis of the resting state . . . . .	106
9.2.1	Electronic and spin structure . . . . .	106
9.2.2	Atomic structure . . . . .	113
9.2.3	Chemical bonds . . . . .	113
9.2.4	Which atom is in the center? . . . . .	115
9.2.5	Hyperfine parameters . . . . .	117
<b>10</b>	<b>Protonation</b>	<b>119</b>
10.1	Protonation sites . . . . .	119
10.1.1	Protonation and monodentate homocitrate . . . . .	120
10.1.2	Protonation and H <sub>2</sub> production . . . . .	122
10.2	The chemical potential of protons . . . . .	123
10.2.1	Protonation state of the resting state . . . . .	125
10.2.2	Protonation state of the first reduced state . . . . .	125
10.2.3	Determination of the chemical potential of protons . . . . .	128
10.3	The chemical potential of hydrogen atoms . . . . .	128
10.3.1	Evaluation of protonation energies . . . . .	128
10.3.2	NH <sub>4</sub> <sup>+</sup> as model for the proton source . . . . .	130
10.4	Consequences on the reaction mechanism . . . . .	131
<b>11</b>	<b>Nitrogen reduction</b>	<b>133</b>
11.1	Nitrogen binding . . . . .	133
11.1.1	Nitrogen binding modes . . . . .	134
11.1.2	Axial nitrogen binding ( <i>A0</i> ) . . . . .	140
11.1.3	Bridged nitrogen binding ( <i>A1</i> ) . . . . .	141
11.1.4	Embedding of bound nitrogen in the protein . . . . .	143

11.1.5	Dependence of the binding energy on the oxidation and protonation state of the cofactor . . . . .	145
11.2	The reaction branches . . . . .	145
11.2.1	Branch 1: Protonation of bridged dinitrogen ( <i>A1</i> ) . . . . .	147
11.2.2	Branch 2: Protonation of axial dinitrogen ( <i>A0</i> ) . . . . .	151
11.2.3	Branch 3: Protonation of the central ligand . . . . .	153
11.2.4	Branch 4: From <i>C0</i> to ammonia release . . . . .	155
11.2.5	Molybdenum as N <sub>2</sub> reduction site . . . . .	159
11.3	Discussion . . . . .	161
11.3.1	Required properties of the cofactor . . . . .	161
11.3.2	Binding site . . . . .	162
11.3.3	Dependence of the results on the choice of the hydrogen chemical potential . . . . .	163
11.3.4	The role of H <sub>2</sub> . . . . .	164
11.4	Conclusion . . . . .	164
<b>12</b>	<b>Nitrogenase and Acetylene</b>	<b>167</b>
12.1	Introduction . . . . .	167
12.2	Computational Details . . . . .	168
12.3	Results . . . . .	168
12.3.1	Acetylene binding modes . . . . .	168
12.3.2	Barriers during acetylene binding . . . . .	170
12.3.3	Protonation . . . . .	171
12.3.4	C <sub>2</sub> H <sub>4</sub> Production . . . . .	171
12.3.5	Acetylene binding energies at different reduction states of the cofactor . . . . .	173
12.4	Discussion . . . . .	173
12.5	Conclusion . . . . .	176
<b>13</b>	<b>Summary</b>	<b>179</b>
	<b>Bibliography</b>	<b>187</b>
<b>A</b>	<b>Appendix</b>	<b>207</b>
A.1	Total spin of a Slater determinant . . . . .	207
A.2	Total energies of intermediates . . . . .	210
<b>B</b>	<b>CV of Johannes Kästner</b>	<b>211</b>
<b>C</b>	<b>Papers</b>	<b>215</b>



## Acknowledgment

A journey is easier when you travel together. Interdependence is certainly more valuable than independence. This thesis is the result of three years of work whereby I have been accompanied and supported by many people. It is a pleasant aspect that I have now the opportunity to express my gratitude to all of them.

First, I thank my advisor, Prof. Peter E. Blöchl, who provided a motivating, enthusiastic, and critical atmosphere during the many discussions we had. Besides of being an excellent advisor, Peter was as close as a relative and a good friend to me.

Many other people have helped me not to get lost during the development of this thesis. I thank the members of my group Jürgen Noffke, Alexander Poddey, Clemens Först, and Axel Ehrich for their continuous support in all concerns, especially in computer-related questions, and for making my life in the institute a lot more enjoyable. Christopher Ashman, Tom Kirchner, Andrei Reyes Huamantincó, Sascha Hemmen, Johanna Däumler, Mike Thieme, and Sven Müller all gave me the feeling of being at home at work.

During workshops and conferences I met many other scientists working on nitrogenase and related fields, and I got into contact with others via mail. Too numerous to mention them all, they helped me a lot in finding my way in this exciting field of research.

Clausthal University of Technology provided me with an ideal working environment which was vital for the research. The “Norddeutsche Verbund für Hoch- und Höchstleistungsrechner (HLRN)” supplied me with huge amounts of computer time, especially during the test phase of their IBM p690 supercomputers. I am also thankful to GWDG for access to their supercomputers previous to HLRN.

I am grateful to my parents, Karl and Monika Schimpl, as well as to my sister Marianne Schimpl. She sometimes managed to take my mind off nitrogenase, for example by carrying me off to Australia and Stromboli. Moreover, I thank my friends Christian Breyer and Angelika Fink.

Most of all, I thank my wife Andrea for supporting me during the last years, for moving to Germany with me and for her love and patience during the PhD period.

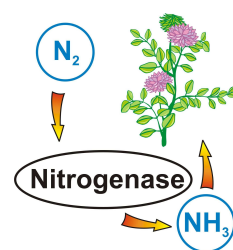


# 1 Introduction

*Science is like sex: sometimes something useful comes out,  
but that is not the reason we are doing it.*

Richard P. Feynman

While nitrogen is abundant in huge amounts in the atmosphere, and is required by plants as fertilizer (nutrient), they can not directly access it. Gaseous nitrogen,  $N_2$ , is bound by one of the strongest covalent bonds in nature. Industry uses the Haber-Bosch process to break this bond and convert dinitrogen to ammonia, requiring about  $500^\circ\text{C}$  and 450 bar. Around  $80 \cdot 10^9$  kg of ammonia are manufactured annually in that way. Nature, in contrast, converts dinitrogen to ammonia at ambient conditions by employing the bacterial enzyme nitrogenase. Biological nitrogen fixation produces two or three times more ammonia per year than the industrial process.



Despite half a century of research on nitrogenase, the mechanism of biological nitrogen fixation on an atomic basis is still unknown. Leigh recalls [1]:

By about 1990, many people, myself included, felt that we only had to know the structure of nitrogenase in order to explain how it actually worked.

The structure became known in 1992. However, the mechanism is still elusive today. Howard and Rees, who have unraveled the structure, stated in 1996 [2]:

A major challenge in the enzymology of nitrogenase is to establish a detailed mechanism for reduction of dinitrogen and other substrates in terms of the structures and properties of the nitrogenase proteins.

There are various reasons, why it took over 10 years with known structure, until reliable models for the mechanism have been proposed. On the one hand, the enzyme is rather unstable and therefore experimentally difficult to access. Up to now it was not possible to characterize intermediates of nitrogen conversion. Moreover, a central ligand in the active site of the protein has only been discovered in 2002.

On the other hand, the system is complex and large which makes it difficult to access it from the theoretical point of view. The active site is a metal-sulfur cluster with a complicated spin structure requiring an elaborate computational model. Christiansen, Dean and Seefeldt summarized in 2001 [3]:

Although there is convincing evidence that substrate binding and reduction occurs at the FeMo-cofactor site of nitrogenase, exactly where and how substrate becomes bound to the FeMo-cofactor and is subsequently reduced during catalysis remains among the most intriguing, and as yet unanswered, questions in bioinorganic chemistry.

As to the complexity of the nitrogenase system, it requires experts from different fields of natural sciences: plant physiology, molecular biology, genetics, biochemistry, inorganic chemistry as well as experimental and theoretical physics. Only working together and gaining insight from each other will make it possible to finally unravel the reaction mechanism.

In the present thesis I present an elaborate model for the nitrogen conversion process on an atomistic level. The model explains most of the available experiments, and does not contradict any relevant data.

The work consists of three parts:

1. An overview of the available experimental as well as theoretical results concerning the enzyme. While I concentrate on the data relevant for the analysis of my results I also briefly discuss mutation studies and substrates alternative to  $N_2$ , which are not directly related to my investigations but useful to get an overview of the system.
2. Part 2 provides an introduction to the applied theoretical approach. Starting from density functional theory (DFT), I discuss the projector augmented wave method defining the basis set in my calculations. I focus on the treatment of the electron spin by means of non-collinear spin density as this is a major improvement of my investigations compared with previous studies. The DFT results can best be understood when related to the Heisenberg model describing the interaction of discrete spins. In this chapter, some previously unpublished results regarding an  $[Fe_3S_4]^+$  cluster are given. This cluster exhibits a non-collinear spin arrangement.
3. The third part shows my main results. The structure of the nitrogenase protein is discussed. No DFT calculations are required to investigate proton paths from the surface of the protein to the active center. After a short summary of the properties of isolated  $N_2$  and its reduction products, I discuss the resting state of the enzyme. By comparing the calculated properties with experiment, it was possible to determine the charge state of the resting

state. This is a prerequisite for reliable studies of the conversion mechanism. Protonations of the active site and the chemical potential of protons and electrons are discussed. Finally, the binding of dinitrogen to the cofactor and its reduction are described. An important feature is the opening of the cluster upon binding of  $\text{N}_2$ . Four possible branches of the conversion of  $\text{N}_2$  to  $\text{NH}_3$  are discussed. After the release of the second ammonia molecule, the catalytic cycle is closed.

Finally, binding and reduction of acetylene at the cofactor of nitrogenase is described. As in the  $\text{N}_2$  reduction mechanism, opening of the cluster is the key to understand the reaction.

My first-principles calculations draw a radically different picture of the mechanism of biological nitrogen fixation than previously envisaged. The findings may lend guidance to the search for new biomimetic catalysts that may equal or even surpass the efficiency of nitrogenase and inspire chemists to design new catalytic systems based on clusters analogous to the FeMo-cofactor of nitrogenase.



## **Part I**

# **Nitrogenase – a review**



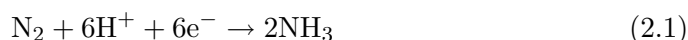


## 2 The protein – structure

*It is wrong to think that the task of physics is to find out how Nature is.  
Physics concerns what we can say about Nature.*  
Niels Bohr

While plants and higher animals are unable to directly use atmospheric nitrogen, special bacteria employ the enzyme nitrogenase for this purpose in a process called biological nitrogen fixation. The most common of these species live in symbiosis with plants as peas, beans, pulses, or lupines of the family of *Leguminosae*.

The general reaction equation of biological nitrogen fixation is



thus dinitrogen ( $\text{N}_2$ ) is reduced to ammonia ( $\text{NH}_3$ ). Various amounts of dihydrogen, not mentioned in the above equation, may be produced as side-products.

Excellent reviews on nitrogenase are available in a thematic volume of Chem. Rev., e.g. [2, 4–6] as well as [7] and some more recent work in [3, 8–10].

### 2.1 Proteins of nitrogenase

Nitrogenase consists of two component proteins named the Fe-protein (or dinitrogenase reductase) and the MoFe-protein (or dinitrogenase). They may occur independently of each other in different concentrations within the cell. The Fe-protein hydrolyzes  $\text{MgATP}$ <sup>1</sup> and uses the required energy to provide electrons to the MoFe-protein. The latter contains the active site of nitrogen fixation, the FeMo-cofactor.

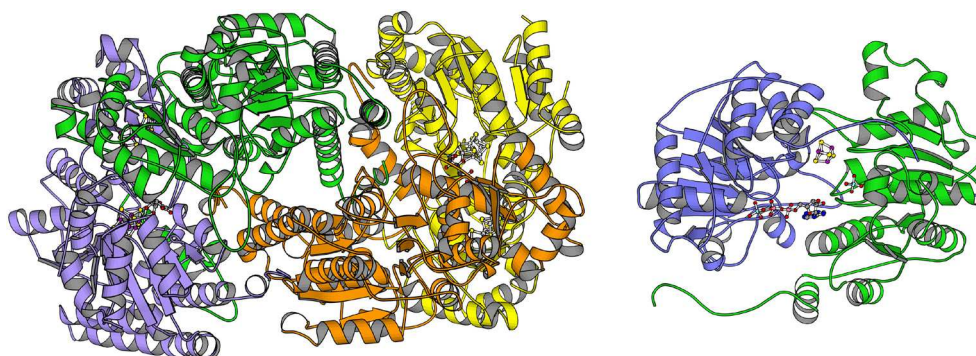
#### 2.1.1 The Fe-protein

The Fe-protein is the smaller of the two nitrogenase proteins, a homo-dimer with a molecular mass of approximately 60-64 kDa<sup>2</sup> [5, 13] containing one  $[\text{Fe}_4\text{S}_4]$  cluster. Its structure was determined in 1992 [12]. A ribbon representation can be seen

---

<sup>1</sup>Adenosine triphosphate, the most important energy transport system in biological matter.

<sup>2</sup>Dalton (Da) is frequently used as the mass unit in biochemistry, 1 Da= $m(^{12}\text{C})/12$ .



**Figure 2.1** Ribbon representations of the MoFe-protein (left) and the Fe-protein (right) from [11,12]. In the MoFe-protein, violet and yellow represent the  $\alpha$  subunits of the  $\alpha_2\beta_2$ -tetramer, while green and orange represent its  $\beta$  subunits. In the Fe-protein, the  $[\text{Fe}_4\text{S}_4]$  cluster as well as one ADP and two  $\text{Mg}^{2+}$  ions are visible. Image source: <http://metallo.scripps.edu/PROMISE/2MIN.html>.

in Fig. 2.1. The functional role is the hydrolysis of two molecules of MgATP per molecule Fe-protein and the subsequent transfer of one electron to the MoFe-protein.

A high concentration of Fe-protein compared to MoFe-protein is called a “high-flux” condition. In this case, electrons (and subsequently also protons) are transferred rather fast to the substrate conversion site.

### 2.1.2 The MoFe-protein

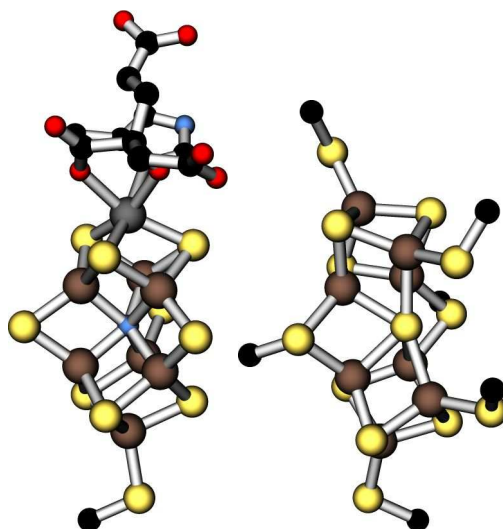
The MoFe protein is an  $\alpha_2\beta_2$ -tetramer containing two  $\alpha$  and two  $\beta$  subunits, with a total mass of approximately 240-250 kDa [5, 13]. Two metal sulfur cluster, called P-clusters, are located at the interface between the  $\alpha$  and  $\beta$  subunits and two FeMo cofactors (M-cluster, FeMoco) are contained within the  $\alpha$  subunits. The FeMo cofactors are assumed to be the location of substrate conversion. Fig. 2.1 shows a ribbon representation of the protein.

The structure was first unraveled in 1992. Since then, increasingly more refined crystallographic structure determinations have been published as summarized in table 2.1. The structure determinations have provided the only reliable information about the stoichiometry of the cofactor. An additional central ligand, which could either be C, N or O, was found in the FeMoco using a very accurate structure determination [13] in 2002. Theoretical investigations [24–27], including my own, represented in section 9.2.4, ruled out oxygen and showed that the most plausible species is nitrogen. Thus the composition of FeMoco is  $\text{MoFe}_7\text{S}_9\text{N}$ .

FeMoco, illustrated in Fig. 2.2, is linked to the protein via a sulfur atom from a cysteine residue (Cys $\alpha$ 275) and a nitrogen atom of the imidazole ring of a histidine

species	protein	refinement	Ref.	PDB	comment
Av	MoFe	2.7 Å	[14, 15]		
Av	Fe	2.9 Å	[12]	1NIP	
Av	MoFe	2.2 Å	[16]		
Av	MoFe	2.0 Å	[11]	3MIN	resting state
Av	MoFe	2.0 Å	[11]	2MIN	oxidized state
Av	Fe	2.2 Å	[17]	2NIP	
Av	Fe	2.25 Å	[18]	1G1M	all-ferrous
Av	MoFe	1.16 Å	[19]	1M1N	central ligand
Av	complex	2.3 Å	[20, 21]	1M34	
Cp	MoFe	3.0 Å	[22]	1MIO	
Cp	Fe	1.93 Å	[17]	1CP2	
Kp	MoFe	1.6 Å	[23]	1QGU	resting state
Kp	MoFe	1.6 Å	[23]	1QH1	oxidized

**Table 2.1** Structures of nitrogenase proteins determined by crystallographic analysis. Abbreviations: Av: *Azotobacter vinelandii*, Cp: *Clostridium pasteurianum* Kp: *Klebsiella pneumoniae*. References to the respective Protein Data Bank (PDB) entries are given.



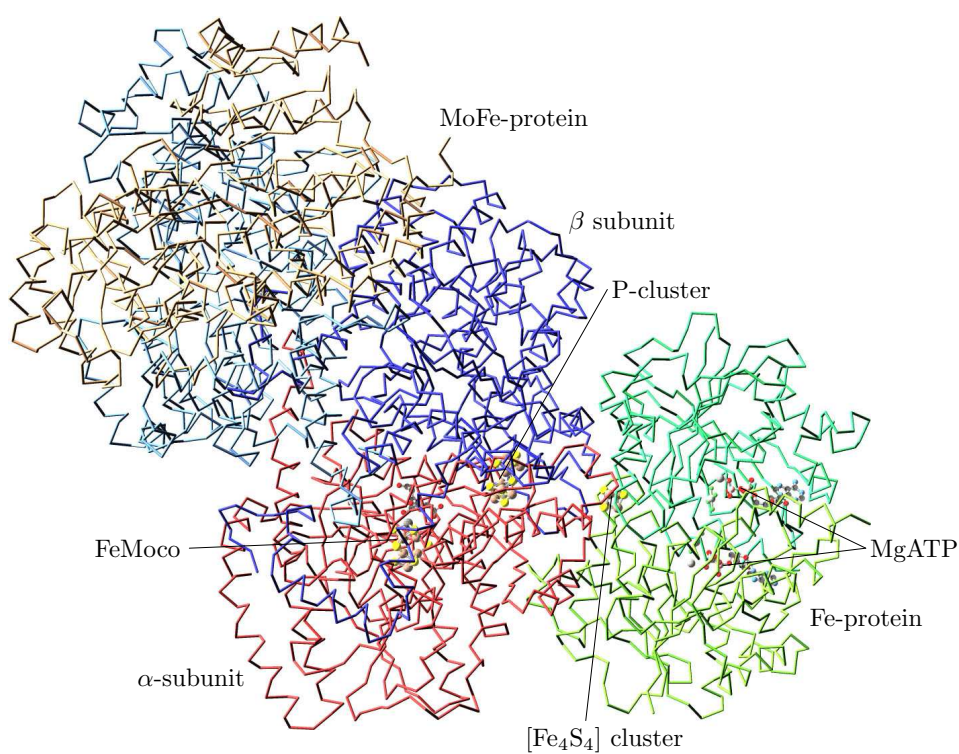
**Figure 2.2** The FeMo-cofactor and its ligands (left) and the P-cluster with the sulfur bonds of cysteine residues (right). Brown: Fe, Yellow: S, Gray: Mo, Black: C, Blue: N, Red: O.

(His $\alpha$ 442). A bidentate homocitrate ligand is connected to the molybdenum atom of the cofactor. All residues in the direct vicinity of the cofactor are conserved in the different nitrogenase bacteria.

However, the sequential numbers of these residues may change due to variations in the protein sequence. Throughout this thesis I will refer to the amino acid sequence of *Azotobacter vinelandii*.

The P-cluster [ $\text{Fe}_8\text{S}_7$ ], shown in Fig. 2.2, enables charge transfer from the Fe-protein to the FeMoco. It has been shown, by crystallographic analysis [11], to undergo structural changes during change of the redox state. The P-cluster is located approximately 14 Å (edge-to-edge distance) from the FeMoco with the region between the P-cluster and the FeMoco largely occupied by the homocitrate ligand and a network of hydrogen-bonded water molecules as discussed in section 7.3 on page 90.

### 2.1.3 Complex formation



**Figure 2.3** Complex between the MoFe-protein (left) and the Fe-protein (right). Only the backbones of the protein chains are shown. The cofactors are only shown in one half of the MoFe-protein. Structure from [21].

The electron transfer from the Fe-protein to the MoFe-protein occurs via formation of a complex between them [5]. The two proteins are arranged such that they minimize the distance between the  $[\text{Fe}_4\text{S}_4]$  cluster of the Fe-protein and the P-cluster in the MoFe-protein. The distance is  $\approx 15 \text{ \AA}$ . A short distance is required to facilitate electron transport. The Fe-protein is only able to hydrolyze MgATP when bound in this complex.

It has been possible to stabilize this complex, crystallize it, and determine its structure [20, 21] by replacing MgATP by MgADP·AlF<sub>4</sub> and thus suppress hydrolyzation. The complex is illustrated in Fig. 2.3 on the facing page.

## 2.2 Cofactors: electronic properties

### 2.2.1 FeMo-cofactor

The major physical property to experimentally identify FeMoco in its resting state is a strong and specific  $S = 3/2$  EPR signal [28]. This signal can be detected in living cells as well as in MoFe-protein isolated from the cells in the dithionite-reduced state, which shows that the charge state is the same in these two conditions.

It is even possible to extract FeMoco from the MoFe-protein while preserving the  $S = 3/2$  signal. The isolated cofactor can reactivate FeMoco-deficient MoFe-protein. This also shows that its basic functions are preserved upon extraction.

The  $S = 3/2$  signal from FeMoco vanishes upon exposure of the isolated cofactor or the protein to molecular oxygen—it only exists under reducing conditions. Structures of the cofactor have been determined in the resting state as well as in a more oxidized state [11]. They only show minor differences.

Three oxidation states of FeMoco are experimentally known and have been characterized to some extent. While the resting state is an  $S = 3/2$  system, the singly oxidized state is diamagnetic,  $S = 0$  [5, 29]. A state reduced by one electron compared to the resting state has an integer spin  $S > 0$  [30]. Even though the relative oxidation states are known, it has not been possible to identify their absolute charge states experimentally.

Mössbauer isomer shifts provide some hints about the oxidation state of the cluster. However, they also depend on the structural environment of the investigated iron site. The mean Mössbauer isomer shifts of the MoFe-protein, with FeMoco selectively enriched by  $^{57}\text{Fe}$ , are 0.41 mm/s, 0.35 mm/s, and 0.43 mm/s for the resting state, the oxidized and the reduced state, respectively [30]. Detailed results of Mössbauer spectroscopy of the resting state are given in table 2.2.

The seven iron atoms of FeMoco are in their high-spin state and thus their hyperfine parameters can be detected with ENDOR and Mössbauer spectroscopy. While only five different iron sites  $A^1, A^2, A^3, B^1, B^2$  could be identified in the ENDOR experiments [31], recent Mössbauer measurements [30] show that the

site	$\delta$ (mm/s)	$\Delta E_Q$ (mm/s)	$\eta$	$A_{iso}$ (MHz)
$A^1$	0.39	-0.69	1	-18.0
$A^2$	0.48	-0.94	1	-17.1
$A^3$	0.39	-0.56	1	-11.8
$A^4$	0.41	0.68	1	-3.7
$B^1$	0.33	-0.66	0.9	11.7
$B^2$	0.50	-0.65	1	9.3

**Table 2.2** Isomer shift ( $\delta$ ), Quadrupole splitting ( $\Delta E_Q$ ), asymmetry parameter ( $\eta$ ), and the isotropic hyperfine coupling parameter  $A_{iso}$  of the resting state of FeMoco determined by Mössbauer spectroscopy [30].

$B^1$  site is doubly degenerate and identify one additional site  $A^4$ . The isotropic hyperfine coupling constants  $A_{iso}$ , indicative of the spin density at the nuclear sites, show that the spins are antiferromagnetically coupled with four negative sites,  $A^1 - A^4$ , and three positive sites,  $2 \times B^1, B^2$ . The site  $A^4$  evaded the first ENDOR studies because of a small hyperfine parameter, indicating that its local spin may be nearly perpendicular to the total spin [30]<sup>3</sup>. Hyperfine parameters and electric field gradients, including their directions, are now available for all iron sites and are given in table 2.2. The hyperfine couplings at the molybdenum atom have been obtained by ENDOR [32]. They are small, showing that the spin density at the Mo site is small.

The structural data from X-ray diffraction are augmented by EXAFS<sup>4</sup> data. They provide information on the bond distances in the unperturbed protein [33–36] or the isolated cofactor [35, 37]. However, EXAFS measurements in the protein always result in a superposition of the bond length of FeMoco and of the P-cluster. The isolated cofactor in the solution may be distorted compared to the one in the protein. Nevertheless the results of EXAFS measurements are used in section 9 to compare experimental and theoretical bond lengths.

### 2.2.2 P-cluster

The P-cluster is known to exhibit a number of charge states. In the dithionite-reduced state  $P^0$ , Mössbauer measurements show that all or nearly all iron sites are in their ferrous ( $Fe^{2+}$ ) state [38, 39]. Thus the charge state of the cluster is not known absolutely, but oxidation states relative to each other are known which rationalize the usage of the nomenclature. The  $P^0$  state is diamagnetic with  $S = 0$ . The various oxidation levels have multiple spin states with facile interconversions [29] and closely spaced energy levels: the  $P^{1+}$  state is a mixture of  $S = 1/2$  and

<sup>3</sup>My calculations do not confirm this finding. They show that the spin alignment of the resting state is collinear, see section 9.2.1 on page 106.

<sup>4</sup>extended X-ray absorption fine structure

$S = 5/2$  contributions [39], the  $P^{2+}$  state is most probably  $S = 3$  or  $S = 4$  [39]. There is also a state  $P^{3+}$  showing  $S = 1/2$  and  $S = 7/2$  as well as a super-oxidized state with  $S > 2$  imposing irreversible damage to the system [5]. It is not yet clear which of these states participate in the catalytic cycle. The P-cluster may play the role of a simple one-electron transfer center or may collect electrons and deliver pairs of them to the FeMoco.

### 2.2.3 $[Fe_4S_4]$ cluster of the Fe-protein

The redox levels which can be adopted by the  $[Fe_4S_4]$  cluster are well known. Within the catalytic cycle the cluster changes between the  $[Fe_4S_4]^+$  and the  $[Fe_4S_4]^{2+}$  level [5, 40]. The reduced state has  $S = 1/2$ . However, a low-lying  $S = 3/2$  state has also been observed [41]. There is one ferrous pair of iron atoms and one mixed-valence pair. In the latter, the spins of two iron sites are parallel and their spin-up shells are filled. Additionally, there is one spin-down electron forming a weak metal-metal bond between them. Thus, both sites are in a formal  $Fe^{2.5+}$  oxidation state.

The oxidized form  $[Fe_4S_4]^{2+}$  consists of two such mixed-valence pairs and is diamagnetic,  $S = 0$  [41].

There is also experimental evidence that the  $[Fe_4S_4]$  cluster can be reduced under biologically accessible conditions to an all-ferrous state  $[Fe_4S_4]^0$  [42, 43] with  $S = 4$  [44]. The structure of such an Fe-protein has been determined [18]. This opens the possibility that under some conditions the Fe-protein may act as a two-electron donor to the MoFe-protein. Implications of this possibility on the reaction mechanism will be discussed in section 3.1. There is also experimental indication that the  $[Fe_4S_4]$  cluster may split into two parts in the oxidized state [45]. These structural changes are induced by MgATP hydrolysis.

## 2.3 Theoretical work concerning the resting state

### 2.3.1 Calculations considering the central ligand

Five independent DFT-studies [24–26, 46] including my own [27], discussed in chapter 9, coincide in their result concerning the oxidation state of the resting state of FeMoco:  $[MoFe_7S_9N]^0$ .

Dance [24] assumed nitrogen to be the central ligand in a study using the BLYP functional and the DMol-program package. He related the energy of the HOMO to the experimentally measurable redox potential of the cofactor and thus found the resting state to be represented by the redox level of  $[MoFe_7N]^{+18}$  which, after inclusion of the sulfur atoms, corresponds to a cluster core-charge of  $[MoFe_7S_9N]^0$ . He also describes a possible mechanism of nitrogen intercalation into the cluster

as well as several N<sub>2</sub> binding modes, although he does not give binding energies. None of his proposed binding modes is equivalent to my results presented in section 11.1 on page 133.

Lovell and coworkers in the group of Noodleman [25] used structure, redox potential and average Mössbauer isomer shifts as criteria for finding the resting state. They use the PB86 [47] functional in the ADF package. C, N, and O are compared as possible candidates for the central ligand. They do not optimize their spin-ordering but use the one obtained from the cluster without the central ligand. It differs from my optimized spin-ordering, discussed in section 9. They conclude it has a resting state of [MoFe<sub>7</sub>S<sub>9</sub>N]<sup>0</sup>.

Hinnemann and Nørskov [46] also found nitrogen to be the central ligand although they used a different charge state than the other groups (functional: RPBE [48], program package: dacapo). Since they are only able to describe an overall neutral system, they used [MoFe<sub>7</sub>S<sub>9</sub>NH<sub>3</sub>]<sup>3+</sup> as resting state.

Vrajmasu et al. in the group of Münck [26] compare calculated mean Mössbauer isomer shifts to experiment and rule out oxygen as the central ligand. Using B3LYP [49] implemented in Gaussian98 with the 6-311G basis set, they simplified the cofactor in order to fit it into *D*<sub>3h</sub>-symmetry. They concluded it has a [MoFe<sub>7</sub>S<sub>9</sub>N]<sup>0</sup> resting state. In an attempt to explain the hyperfine parameters they assumed a non-collinear spin distribution using intrinsic hyperfine parameters from model complexes.

### 2.3.2 Calculations without considering the central ligand

Numerous theoretical studies had been published before the central ligand was discovered [50–75]. They are of limited value now, since the central ligand has significant effects on the chemistry of the cluster. However, here I provide a brief overview of them. All calculations reported up to now are restricted to collinear spin arrangements.

The most detailed study of the resting state of the cofactor without central ligand has been done by Lovell, Li, Case and Noodleman [72]. They used DFT with BP86 [47] functional and uncontracted triple- $\zeta$  STOs as a basis set. In a model with 72 atoms they included the cofactor, the whole homocitrate as well as some H-bonding partners of the latter. Based on experimental ENDOR results [76], they used a charge state with one ferric iron but also investigated the state with three ferric iron sites. All 10 broken-symmetry states, representing an *S* = 3/2 state in 3-fold rotational symmetry with one ferric site, have been calculated. For the state with the lowest energy, they removed the 3-fold symmetry. No clear decision can be made as to which of the 3 resulting states represents the ground state.

They also investigated if the state with three ferric sites could represent the



resting state and performed 12 calculations to test all collinear spin arrangements representing an  $S = 3/2$  state. Geometries as well as the nature of the frontier orbitals change little on the two electron oxidation. No spin is assigned to Mo in either of the charge states, but one unpaired electron resides on the homocitrate ligand.

Rod and Nørskov [62, 64] have used GGA-DFT with the pseudopotential method. They used a model where the core of the cofactor is periodically repeated in one direction, with the Mo replacing the tetrahedrally coordinated Fe-atom of the next unit. Using two ferric iron sites, they obtain a spin  $S = 0$ . In [63] they also used an isolated model with ligands truncated as SH, OH, and  $\text{NH}_3$ . They obtained a spin of  $S = 3/2$  using an oxidation state with three ferric iron sites as the resting state because they have to use charge-neutral cells.

Siegbahn et al. [61] did most of their calculations on smaller model systems with the B3LYP hybrid-functional [49]. The largest system was the core of the cofactor without any ligands, and the molybdenum replaced by iron. They used ferromagnetic coupling. The Fe-Fe' distance decreased significantly upon oxidation.

Dance [52, 54] only investigated two oxidation states with an even electron number, which cannot represent a half-integer spin state.

Stavrev and Zerner [57, 58] used intermediate neglect of differential overlap (INDO) to calculate a model of the whole cofactor with truncated ligands. They calculated models with 1-6 ferric iron sites and found an  $S = 3/2$  ground state in all even charge states, degenerate with the  $S = 1/2$  state in the case of three ferric sites. However, they used the crystal structure refined to 2.2 Å [16] for all charge states and did not relax it. They found that all bond indices ("Bindungsordnung" which is 3 for  $\text{N}_2$ , 2 for  $\text{O}_2$ ) decrease upon reduction except the  $\text{Fe-S}^3$  bond which increases.

The earliest electronic structure calculations [50], based on extended Hückel theory, occurred shortly after publication of the crystal structure. The value of this work is of limited value today because of the neglect of spin polarization.



## 3 The mechanism

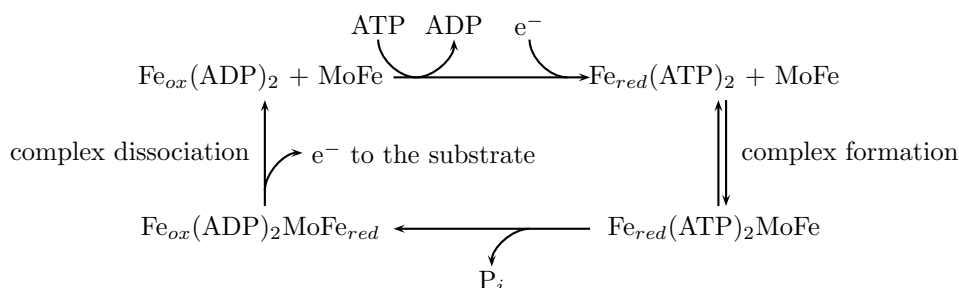
*A common mistake people make when trying to design something completely foolproof is to underestimate the ingenuity of complete fools.*

Douglas Adams

The reaction mechanism of nitrogenase can be separated into two nested cycles: (1) The whole cycle of  $N_2$  being reduced, which requires about eight electrons, depending on the amount of  $H_2$  produced; (2) and the Fe-protein cycle which provides these electrons one-by-one. I start with a discussion of the latter.

### 3.1 The Fe-protein cycle

The Fe-protein hydrolyzes MgATP and provides electrons to the MoFe-protein. I describe this cycle starting from the oxidized form of the protein bound to two molecules of MgADP<sup>1</sup>, denoted  $Fe_{ox}(MgADP)_2 + MoFe$ . In this stage the Fe-protein is not bound to the MoFe-protein.



**Figure 3.1** The Fe-protein cycle

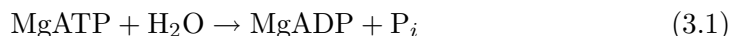
*In vitro*, the Fe-protein will be reduced by dithionite, the usual reducing agent used in these experiments. In the cell, *in vivo*, the reduction will be done most probably by electron carriers like for example  $NAD^+$ ,  $NADP^+$ , or  $FAD^2$ , obtaining

<sup>1</sup>adenosine diphosphate

<sup>2</sup>nicotinamide adenine dinucleotide, nicotinamide adenine dinucleotide phosphate, flavine-adenine dinucleotide

their electrons from the citric acid cycle or from glycolysis. This reduction is accompanied by a replacement of the two MgADP molecules by new MgATP molecules, leading to  $\text{Fe}_{red}(\text{MgATP})_2 + \text{MoFe}$ , see Fig. 3.1.

The next step is the complex formation between the proteins which requires the Fe-protein to be in its reduced state and bound to MgATP. Contraction of the proteins enables complex formation [5], which is reversible, leading to  $\text{Fe}_{red}(\text{MgATP})_2\text{MoFe}$ .



In this complex, MgATP is cleaved to MgADP and phosphate ( $\text{P}_i$ )<sup>3</sup> as shown above. One electron is transferred to the MoFe-protein. The order of these two processes is uncertain. They may occur simultaneously [5]. One phosphate from MgATP is released. It leads to  $\text{Fe}_{ox}(\text{MgADP})_2\text{MoFe}_{red}$ .

Next, the complex dissociates, which is believed to be the rate-limiting step not only for the Fe-protein cycle but for the whole substrate conversion. The rate of complex dissociation is about  $5\text{ s}^{-1}$  [77] under saturating conditions. It leads to  $\text{Fe}_{ox}(\text{MgADP})_2 + \text{MoFe}_{red}$ .

The electron at the MoFe-protein is used for substrate conversion and the Fe-protein is now again in its starting configuration closing the Fe-protein cycle.

There is experimental evidence that the  $[\text{Fe}_4\text{S}_4]$  cluster exists at three different oxidation states at biological accessible reduction potentials [42, 43]. There is, however, no evidence that all three of them take part in the Fe-protein cycle. Stopped flow measurements, described in the next section, show that the kinetics of the reaction can be well described with the assumption that only one electron is transferred in each Fe-protein cycle.

The rate of the Fe-protein cycle is thus limited by the complex dissociation, which also limits the electron transport to the active site of nitrogenase.

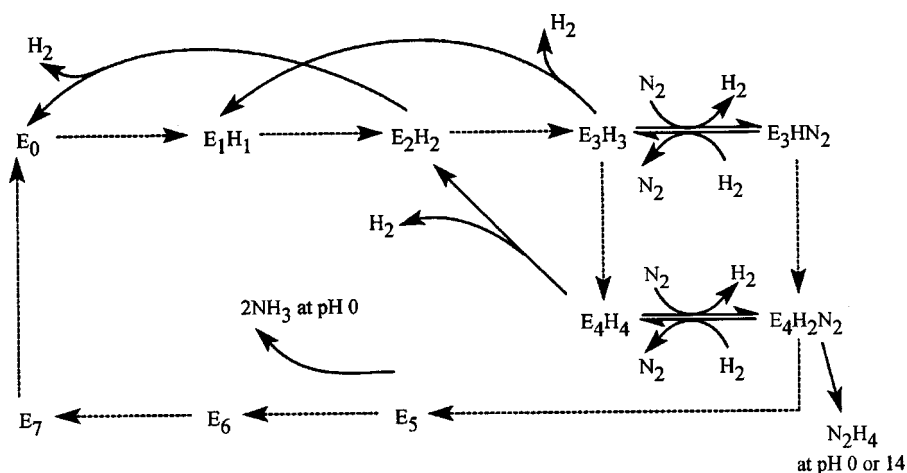
## 3.2 The Thorneley-Lowe scheme for nitrogen fixation

In 1984 Thorneley and Lowe [78–80] presented a scheme based on kinetic stopped-flow measurements describing the nitrogenase  $\text{N}_2$  turnover cycle. Its main conclusions are that

- the cycle is controlled by the electron donation of the Fe-protein cycle,
- protons and electrons are added one-by-one to the MoFe-protein,
- $\text{N}_2$  binds after the protein has been reduced by 3-4 electrons.

---

<sup>3</sup> $\text{P}_i$  an abbreviation for phosphate in an equilibrium between  $\text{HPO}_4^{2-}$  and  $\text{H}_2\text{PO}_4^-$



Starting from the resting state  $E_0$ , the MoFe-protein is reduced by one electron and protonated once, leading to state  $E_1H_1$  as depicted in Fig. 3.2. The notation  $E_xH_y$  describes the protein with  $x$  electrons and  $y$  protons added to the resting state.

$\text{E}_2\text{H}_2$  will again be reduced and protonated, leading to  $\text{E}_3\text{H}_3$ , which again can produce  $\text{H}_2$ . In this state dinitrogen can bind. This produces one  $\text{H}_2$  and leads to  $\text{E}_3\text{HN}_2$ . In its turn,  $\text{H}_2$  can displace  $\text{N}_2$  and thus inhibit its reduction.  $\text{N}_2$  is the only nitrogenase substrate whose reduction is inhibited by  $\text{H}_2$ .

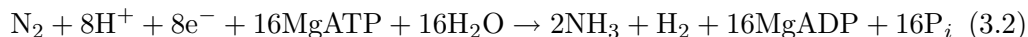
It is known from kinetics that substrates bind only to, and products are only released from, free MoFe protein not bound in the complex to the Fe-protein. Additionally it is known that the site(s), at which the substrates enter and the products leave the protein, are protected from access by solvent in the long living complex [5]. Since the FeMo-protein undergoes its major changes at the binding region of the Fe-protein during complex formation, it has been suggested that the substrates enter and leave via the homocitrate – P-cluster pathway, which is

described in [9].

It is known from EPR experiments that the  $S = 3/2$  signal from FeMoco vanishes during turnover. However, some intermediates are also known to exhibit an  $S = 3/2$  signal [82]. EPR measurements during turnover show that only two of the first three electrons initially transferred to the protein actually reach the FeMo-cofactor [82]. Thus the reduction states of the Thorneley-Lowe scheme must not be directly mapped onto reductions of FeMoco.

The Thorneley-Lowe scheme suggests a minimum amount of one mole of  $H_2$  produced per mole of converted  $N_2$ . While, in general, the enzyme will work less efficiently, producing more  $H_2$ , there is also experimental evidence [83] that nitrogenase may act more efficiently under optimal conditions. A complete suppression of  $H_2$  production, however, can also be ruled out, even under high pressure of  $N_2$  [84].

The conversion of  $N_2$  to ammonia requires six electrons, thus consumes 12 units of MgATP. Each sacrificially produced mole of hydrogen additionally consumes 4 moles of MgATP. Thus the general reaction equation of nitrogenase turnover may be written as



with MgATP being hydrolyzed to MgADP and phosphate ( $P_i$ ). In pH-neutral conditions most of the ammonia will be protonated to  $NH_4^+$ .

The total reaction rate depends on reaction conditions and on the species the enzyme has been obtained from. Under optimal conditions it is around 1204 nmol  $NH_3$  formed per minute per milligram protein of *Azotobacter vinelandii* [85], which corresponds to about one  $N_2$  conversion per second per molecule of FeMoco.

### 3.2.1 Nitrogen and hydrogen

There are three types of interaction of nitrogenase with molecular hydrogen:

1.  $H_2$  is always produced during  $N_2$  turnover [86], but the ratio between  $N_2$  and  $H_2$  does not seem to be fixed in neither direction. This is called OHE (ordinary hydrogen evolution).  $H_2$  is released when  $N_2$  binds to the cluster [79] and can on the other hand displace bound  $N_2$ .  $H_2$  is a competitive inhibitor<sup>4</sup> of  $N_2$  turnover, but of no other substrate [87].
2.  $H_2$  is produced by nitrogenase if no other substrate is available [88].

---

<sup>4</sup>Competitive inhibitors compete with normal substrate for the active site of an enzyme while non-competitive inhibitors act on other parts of the enzyme or on the cofactor and decrease the enzyme's ability to combine with substrate.

reaction	$\Delta E'^{\circ}$	
$N_2 + 8H^+ + 6e^- \rightleftharpoons 2 NH_4^+$	-0.28 V	[10]
$2H^+ + 2e^- \rightleftharpoons H_2$	-0.421 V	[91]

**Table 3.1** Standard reduction potentials at pH= 7.

3. HD formation [89]: in the presence of  $N_2$ , and of no other substrate,  $D_2$  gas is converted to HD. No HD formation takes place at the limit of infinite partial pressure of  $N_2$  [90].

It has been concluded from experiments on the reaction kinetics [5] that  $H_2$  inhibition of  $N_2$  turnover and HD formation may be manifestations of the same underlying process. HD reduction (turnover) does not produce  $D_2$ .

If  $T_2$  is offered, HT is produced, but no tritium in the form of  $T^+$  or HOT can be found in the water surrounding the protein. Thus the atoms from molecular hydrogen never dissociate from the cofactor except in the form of molecular hydrogen.

This finding can be interpreted as dissociation of the  $H_2$  molecule at the cofactor, with the products tightly bound to the cofactor in the form of, for example,  $H^-$ . These tightly bound atoms may react with protons from the solvent and leave the complex as molecular hydrogen.

From the thermochemical point of view, it is preferable for any enzyme to produce hydrogen instead of reducing nitrogen to ammonia. The free-enthalpy difference is 11.6 kJ/mol per  $H_2$ , which can be calculated from the redox potentials given in table 3.1. Thus nitrogenase has to suppress  $H_2$  production kinetically. It has been suggested that this is done by slow, controlled proton transport to the cofactor. The channel responsible for this transport is discussed in section 7.1.3.

### 3.2.2 Other substrates and inhibitors

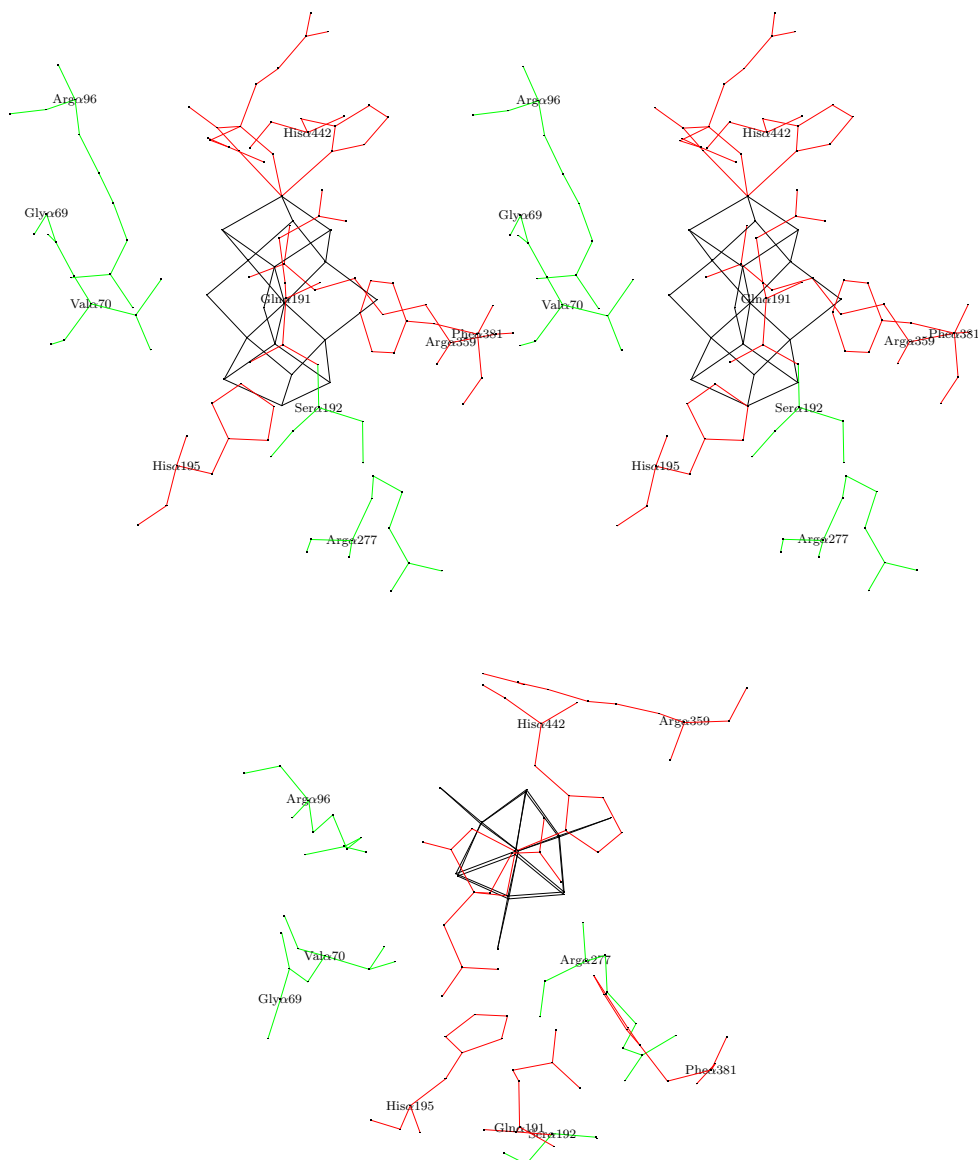
The most important natural substrates of nitrogenase are nitrogen and protons. The latter are reduced to  $H_2$  if no other substrate is available.

Besides those there is a long list of known substrates, which vary in size from the proton up to  $CH_2=CHNC$  and 1-butyne. Known substrates are [5]:  $N_2$ ,  $CH_3NC$ ,  $CN^-$ ,  $N\equiv CNH_2$ ,  $C_2H_2$ ,  $C_2H_4$ ,  $N_2O$ ,  $COS$ ,  $CO_2$ ,  $CO$ ,  $H_2$ ,  $H^+$ ,  $N_2H_2$ ,  $N_3^-$ ,  $CH_3C\equiv CH$ ,  $C_2H_5C\equiv CH$ ,  $H_2C=C=CH_2$ ,  $CH_2N=N$ ,  $C_2H_5CN$ ,  $C_3H_7CN$ ,  $C_2H_5NC$ ,  $H_2C=CHNC$ , *cis*- and *trans*- $CH_3-N=N-CH_3$ . Some of them are reduced, some, like for example  $CO$ , only bind to the cofactor, but are not reduced.

residue	replaced by	reduces N <sub>2</sub> ?	ref.	comment
MoFe protein $\alpha$ -chain				
Gly69	Ser	+	[92]	no C <sub>2</sub> H <sub>2</sub> reduction
Val70	Ala	+	[13, 93]	reduces short chain alkynes
Arg96	Gln	+	[8, 94]	binds C <sub>2</sub> H <sub>2</sub> and CN <sup>-</sup> at the resting state
Arg96	Leu, Ala, His	?	[94]	binds C <sub>2</sub> H <sub>2</sub> and CN <sup>-</sup> at the resting state
Arg96	Lys	+	[85]	
Gln191	Lys	-		
Ser192	Asn, Thr, Gly, Leu, Val	+	[95]	
His195	Gln	$\pm$	[95]	Structure determinated [95].
His195	Asp	-	[5, 95]	
His195	Tyr, Leu, Thr, Gly	-	[96, 97]	
Arg277	Lys, Cys, Thr, Phe, Leu	+	[98]	
Arg277	His	-	[8, 98]	
Arg359	Lys	+	[85]	
Arg359	Gln	-	[8, 85]	no $S = 3/2$ signal
Phe381	Leu,Ile	+	[85]	
Phe381	Arg	-	[8]	
His442	Asn, Cys	-	[85]	no $S = 3/2$ signal
HCA	citrate	-	[99, 100]	

**Table 3.2** Substitutions of amino acids in the MoFe-protein in the vicinity of the cofactor. The ability of the respective protein to reduce N<sub>2</sub> is given.





**Figure 3.3** The residues around FeMoco of which mutation experiments have been performed. Whenever the replacement of an individual residue still enables azeotropic growth it is drawn in green, otherwise in red. Top: stereo view from the front. Bottom: view from the top.

### 3.3 Mutagenesis experiments

Mutagenesis, the replacement of certain amino acids by others, allows getting insight into the significance of individual protein residues on substrate reduction. It can be tested if mutated cells are still able to reduce  $N_2$  (azeotropic growth) or any other substrate. Table 3.2 on page 22 summarizes these results. Those residues near the cofactor, which have been investigated in mutagenesis experiments, are shown<sup>5</sup> in Fig. 3.3 on the page before. The effect of their replacement on azeotropic growth is indicated.

Here, a more detailed discussion of single amino acid replacements is given:

- Gly69 replaced by serine still enables reduction of nitrogen, but suppresses the usual ability of the enzyme to reduce  $C_2H_2$  [92].
- Val70 replaced by alanine still enables azeotropic growth. Even more, it also enables the enzyme to reduce short chain alkynes (like propyne,  $CH_3-C\equiv CH$ ) and propargyl alcohol which are not effectively reduced by the wild type enzyme. This shows that short chain alkynes can bind and be reduced at a specific 4Fe-4S face of the cofactor [13].
- Arg96 replaced by glutamine slows down  $N_2$  reduction [8]. Replacement by leucine, glutamine, alanine, or histidine makes it possible to bind  $C_2H_2$  or  $CN^-$  to the cofactor in the resting state [94] as detected by the appearance of a new EPR signal. However, this binding is reversible: if  $C_2H_2$  or  $CN^-$  are removed from the solution, the new EPR signal vanishes. CO did not prevent  $C_2H_2$  from being bound to the cofactor. CO or  $N_2$  do not change the cofactor EPR signal, i.e. are not bound at the resting state. A  $^{13}C$  hyperfine signal is detected by ENDOR for  $CN^-$  [94]. Arg96 may also be replaced by lysine without suppressing azeotropic growth [85].
- Gln191 replaced by lysine prevents  $N_2$  reduction and affects acetylene reduction.
- Ser192 (not directly at the cofactor) may be replaced by asparagine, threonine, glycine, leucine, and valine [95] without suppressing azeotropic growth.
- His195 may be replaced by glutamine, but not by asparagine without suppressing azeotropic growth. When replaced by glutamine, nitrogenase cannot reduce  $N_2$  very well ( $< 2\%$  of the wild type) but it still induces HD-formation. Therefore  $N_2$  must bind in the same way as in the wild-type

---

<sup>5</sup>In this figure, a stereo view is shown. All such stereo views throughout this thesis have to be viewed as follows: looking at them from a distance of typically 30 cm, the left eye has to focus on the right picture while the right eye has to focus on the left picture. If the pictures of both eyes merge, the structure can be seen three-dimensional.

enzyme. Bound  $N_2$  inhibits acetylene or proton reduction [97]. The structure of the His195Gln mutant has been refined to 2.5 Å by X-ray diffraction in [95].

When His195 is replaced by asparagine,  $N_2$  is bound but not reduced. It still inhibits the reduction of both protons and  $C_2H_2$  but does not induce HD formation [101]. Added  $H_2$  relieves  $N_2$  inhibition of substrate reduction.

His195 replaced by tyrosine, leucine, threonine, or glycine leads to proteins unable to reduce  $N_2$  [96, 97, 102].

- Arg277 may not be replaced by histidine [8, 98] while it may be replaced by lysine, cysteine, threonine, phenylalanine, or leucine without loss of  $N_2$  reduction ability, although these result in diminished reduction rates [98].
- Arg359 replaced by glutamine removes activity and even the  $S = 3/2$  EPR signal of the cofactor [8, 85]. However, replacement by lysine still enables azeotropic growth [85].
- Phe381 replaced by arginine [8] results in loss of azeotropic growth, while its replacement by the neutral residues leucine and isoleucine preserves the ability to reduce  $N_2$  [85].
- His442 replaced by either asparagine or cysteine results in a protein showing no  $S = 3/2$  EPR signal of the cofactor and no catalytic activity [85]. This shows the relevance of the histidine residue connected covalently to Mo.
- The NifV<sup>-</sup> mutant exhibits citrate instead of homocitrate. It is still able to reduce protons but its ability to reduce  $N_2$  is only 7% of the wild type enzyme. The corresponding X-ray structure has been reported in [100].

If  $C_2H_2$  is reduced by the His195Gln mutant cofactor, at least 2  $C_2H_2$  adducts are bound to one cofactor site [95].

Besides mutants there are also alternative nitrogenases in which Mo is substituted by V or Fe. Both are still able to reduce nitrogen, although their efficiency is reduced. There is also a nitrogenase with Mo substituted by tungsten (W) [103]. It does not reduce  $N_2$  and is only marginally active in acetylene reduction, although it still produces  $H_2$  from protons.

### 3.4 Theoretical work concerning the mechanism

The main goal of theoretical investigations on nitrogenase is to unravel its catalytic mechanism. Investigations of the resting state, discussed in section 2.3, provide its charge state, a necessary ingredient for reliable analysis of the mechanism.

### 3.4.1 Calculations considering the central ligand

Only two studies on the reaction mechanism which explicitly considered the central ligand have appeared so far. Hinnemann and Nørksov [104] recalculated their mechanism proposed from results of the vacant cofactor [64] and found it to be plausible also with the central ligand. In their model, the cofactor stays rather rigid during the whole cycle and none of its bonds is broken.  $\text{N}_2$  connects to one of the prismatic iron sites and is reduced step-by-step there. Hydrazine, bound to one Fe site, is an intermediate which explains the hydrazine production of nitrogenase under acid or base quench. After addition of one proton to hydrazine, ammonia desorbs. The authors did not investigate any transition state energies, except the barrier for nitrogen intercalation into the vacant center. For this intercalation they found barriers too high to overcome.

A recent proposal [105], based on the BP86 [47] DFT functional and using the turbomole program package, suggests an opening of a sulfur bridge upon coordination of water to an Fe-site. After a complete protonation of the central ligand, it dissociates as ammonia, and dinitrogen binds in the meanwhile vacant central cavity. Subsequently, one nitrogen atom is fully protonated and dissociated, which closes the catalytic cycle. This intriguing proposal, however, seems to be in conflict with isotope exchange (ESEEM/ENDOR) experiments [106] that exclude an exchange of a central nitrogen ligand.

Durrant [67–70] performed DFT studies on small organometallic model complexes with the molybdenum site and 0–1 iron sites. While these models may represent the cofactor with or without central ligand, I feel that truncation errors may affect the results. Durrant proposes that  $\text{N}_2$  binds to molybdenum. After one protonation it changes into a bridging position between Mo and Fe. The further reduction steps take place at the Mo site. The whole mechanism has been worked out in [69].  $\text{N}_2$  reduction is related to hydrogen evolution also taking place at the Mo site, which is discussed in [70]. Durrant et al. [107] also use their small model to investigate  $\text{CN}^-$  binding to the cofactor.

A prerequisite for nitrogen adsorption at the Mo site is that homocitrate becomes monodentate, i.e. loses one bond to the Mo site and thus opens a coordination site [108]. Grönberg et al. [99] suggest that a hydrogen bond from the longer arm of homocitrate to N $\epsilon$  of His $\alpha$ 442 may stabilize a monodentate coordination of homocitrate.

One interesting idea in [69] is that more  $\text{N}_2$  could be weakly bound to the Fe atoms before the reaction starts at the Mo site, in order to increase the  $\text{N}_2$  concentration. The solubility of dinitrogen in water is only about 0.7 mmol/l.

### 3.4.2 Calculations without consideration of the central ligand

Dance [52–55] performed BLYP DFT calculations and investigated 9 binding modes. The most stable binding site is  $\eta_1$  binding to one of the six under-coordinated Fe atoms, which draws the Fe atom out of the core. In a less stable  $\eta_2$  structure, one nitrogen atom rests on top of the center of a square formed by four under-coordinated iron atoms, while the other nitrogen atom tilts towards its sides. Here the nitrogen bond is expanded from 1.1 Å to 1.29 Å, indicative of a substantial weakening. The proximity of the bridging sulfur atoms (2.7–3 Å) increases the proton affinity of nitrogen, which suggests a proton transfer to the distal nitrogen atom. In the proposed mechanism the distal nitrogen is fully protonated resulting in ammonia formation, leaving behind a nitrogen atom. However, Dance has also investigated a mode with N<sub>2</sub> partially in-plane of a face spanned by four Fe sites [54].

Rod and Nørskov [62–64, 109] suggested a similar mechanism as Dance. Their study is guided by the observation that, on catalyst surfaces such as Fe and Ru, the bond breaking of N<sub>2</sub> is not the rate limiting step. The rate is instead determined by the strong bond between the metal and the nitrogen atoms [110–112]. They use spin-polarized gradient corrected DFT calculations and a chain of connected and periodically repeated clusters MoFe<sub>6</sub>S<sub>9</sub>. The catalytic process begins with  $\eta_1$  binding of N<sub>2</sub> to one three-fold coordinated iron atom, as identified by Dance [55] as the most stable site. The adsorption is followed by alternate hydrogenation and reduction steps leading to N<sub>2</sub>H<sub>4</sub>. Hydrogen is added alternating to one and the other nitrogen. Addition of a further hydrogen to N<sub>2</sub>H<sub>4</sub> results in bond cleavage and production of ammonia. Finally, the NH<sub>2</sub> is converted to the second ammonia. Their conclusion is that the efficiency of nitrogenase is due to the high chemical potentials of the proton and electron reservoirs.

CNDO calculations have been performed by Zhong and Liu [60]. These calculations indicate that, in the most likely binding mode, one nitrogen atom is fully immersed into the center of the Fe cage.

N<sub>2</sub> reduction at Mo, similar to the model of Durrant, has been investigated by Grönberg et al. [99] and Szilagyi et al. [66]. The latter use B3LYP (Gaussian98) calculations on smaller Mo-based models. They regard Mo as being in an  $S = 1$  spin state (two unpaired electrons). A spontaneous bidentate—monodentate transformation of homocitrate has not been found.

A DFT study has been performed by Lovell and coworkers [73] who investigated addition modes of hydrogen to the cofactor with vacant center.



## 4 Open questions

After giving an overview of the facts known about nitrogenase I, will now summarize the questions still open, which can be addressed by DFT calculations and which will be answered in this thesis.

**Charge state of the resting state.** One of the most significant measurable quantities of the cofactor is its  $S = 3/2$  spin signal. This shows that a well-defined, odd number of electrons always resides on the cluster. This defines its charge state.

The chemistry of the cofactor changes upon reduction and protonation. Therefore reliable information about the charge state of the resting state is essential for qualitatively correct simulations of the mechanism.

**Atom type of the central ligand.** Crystallographic analysis has restricted the central ligand to be C, N, or O. While there are some indications that it is nitrogen, its identity has not been clearly determined up to now. The different sizes of these atom types change the reactivity of the central cage.

**Binding modes of  $N_2$ .** At the heart of the open questions is that of the binding modes of dinitrogen to FeMoco. They are the starting points for the reduction mechanism. The protonation states have to be investigated before answering the question of  $N_2$  binding modes.

**The catalytic mechanism.** This is the main goal of this thesis. Unraveling the catalytic mechanism of nitrogenase is the key to understand this enzyme. It may help designing model complexes as well as novel multi-center catalysts, able to break strong bonds.





# **Part II**

# **Theory**



## 5 Electronic structure calculations

*Es gibt nichts Praktischeres als eine gute Theorie.*

Immanuel Kant

Accurate calculations of molecular and solid-state properties require detailed description of the electrons because they are responsible for chemical bonds. In a quantum mechanical description, both electronic and nuclear properties are contained in one many-particle wave function. The so-called Born-Oppenheimer approximation allows to separate the electronic and the nuclear wavefunctions. The nuclei can be treated classically as their mass is at least three orders of magnitude larger than the electronic mass. This mass difference causes a difference in the equilibrium velocity of the particles: electrons move considerably faster than nuclei and can thus be considered as reacting instantaneous on changes of the atomic positions. Thus the electrons are considered to be in their ground state in each atomic configuration. In the following, I only consider the energy of the electronic system at a given atomic geometry.

The properties of such a quantum many-electron system are described by the many-particle Schrödinger equation:

$$\mathcal{H}\Psi = E\Psi \quad (5.1)$$

with the Hamiltonian

$$\mathcal{H} = \sum_{i=1}^N -\frac{\hbar^2}{2m_e} \nabla_i^2 + \sum_{i=1}^N v_{ext}(\vec{r}_i) + \frac{e^2}{4\pi\epsilon_0} \sum_{i<j}^N \frac{1}{r_{ij}} \quad (5.2)$$

where  $\Psi$  is the many-electron wave function,  $E$  is the total energy,  $\nabla_i^2$  is the Laplacian with respect to the coordinates of electron  $i$ ,  $v_{ext}$  is the external (Coulomb) potential of the nuclei and  $r_{ij}$  is the distance between the electrons  $i$  and  $j$ . The energy of the system is given by the variational principle, minimizing the functional

$$E[\Psi] = \frac{\langle \Psi | \mathcal{H} | \Psi \rangle}{\langle \Psi | \Psi \rangle}. \quad (5.3)$$

While Schrödinger's equation for the hydrogen atom and some other simple one-particle problems can be solved analytically, no such solution exists for complicated

many-body systems. However, in principle it is possible to solve these problems numerically using multi-reference configuration interaction or quantum Monte-Carlo methods. These methods are computationally very demanding and thus only allow to treat small systems.

## 5.1 Density functional theory

There is a possibility for mapping the interacting many-particle system onto many non-interacting effective one-particle systems: density functional theory.

### 5.1.1 Hohenberg-Kohn theorems

The Hohenberg-Kohn theorems allow to completely characterize the Hamiltonian of a system from its electron density. Additionally, an energy variational principle for the density is provided.

For any electronic system described by the Hamiltonian (5.2) both the ground state energy and the ground state wave function are determined by the minimization of the energy functional  $E[\Psi]$  of equation (5.3). The external potential  $v_{ext}(\vec{r})$  and the number of electrons  $N$  completely determine the Hamiltonian and thus all electronic properties of the system.

In place of  $N$  and  $v_{ext}(\vec{r})$ , the first Hohenberg-Kohn theorem [113] allows the use of the electron density<sup>1</sup>  $n(\vec{r})$  as basic variable:

The external potential  $v_{ext}(\vec{r})$  is determined, within a trivial additive constant, by the electron density  $n(\vec{r})$ . [113]

For a proof using only the minimum-energy principle, see the book by Parr and Yang [114]. Thus, the electron density determines both  $v_{ext}(\vec{r})$  and  $N$ . Therefore  $n(\vec{r})$  is sufficient to determine the Hamiltonian, the many-particle ground state wave function  $\Psi$ , and thus all other electronic properties.

The kinetic energy  $T(n)$ , the potential energy  $V(n)$  with the part of nuclei-electron interaction  $V_{ne}$  and the part of electron-electron interaction  $V_{ee}$ , as well as the total energy  $E(n)$  may be expressed as functionals of  $n$ :

$$E[n] = T[n] + V_{ne}[n] + V_{ee}[n] \quad (5.4)$$

$$= \int n(\vec{r})v_{ext}(\vec{r}) d^3r + F[n] \quad (5.5)$$

Where  $F[n] = T[n] + V_{ee}[n]$  is a universal functional, independent of  $v_{ext}(\vec{r})$  and thus valid for all systems. Unfortunately, the exact form of this functional is unknown. Approximations will be discussed in sections 5.1.3 and 5.1.4.

---

<sup>1</sup>The charge density  $\rho$  can be obtained from the electron density  $n$  as  $\rho = -en$ .

The second Hohenberg-Kohn theorem provides an energy variational principle for the density analogous to the one for the wave functions (5.3):

For a trial density  $n'(\vec{r})$ , such that  $n'(\vec{r}) \geq 0$  and  $\int n'(\vec{r}) d^3r = N$ ,  $E_0 \leq E[n']$  where  $E[n']$  is the energy functional of equation (5.5) with the external potential  $v_{ext}(\vec{r})$  of the system. [113]

### 5.1.2 Kohn-Sham method

The Hohenberg-Kohn theorems provide an exact method of calculating all electronic properties of a system out of the electron density  $n$ , a quantity depending on the three spatial coordinates. This is an enormous simplification compared to the many-particle wavefunction  $\Psi$  depending on  $3N$  coordinates<sup>2</sup>. However, both parts  $T[n]$  and  $V_{ee}[n]$  of the density functional  $F[n]$  are unknown. Only crude approximations, like the Thomas-Fermi method and related models, described in [114], are available for these functionals.

Kohn and Sham [115] addressed this problem by invoking a *non-interacting reference system* to the true many-particle system with the total density

$$n(\vec{r}) = \sum_i^N \sum_s |\psi_i(\vec{r}, s)|^2. \quad (5.6)$$

Any non-negative, continuous, and normalized density  $n$  can be decomposed into one-particle wave functions  $\psi_i$  according to equation (5.6)<sup>3</sup>. To find a unique set of  $\psi_i$  they are constructed to represent a reference system with the same ground state density as the true system. As the reference particles do not interact, the wave function is of one-determinant manner:

$$\Psi_s = \frac{1}{\sqrt{N!}} \det[\psi_1 \psi_2 \cdots \psi_N] \quad (5.7)$$

The kinetic energy  $T_s[n]$  of this non-interacting system can thus be calculated exactly and is expected to be close to the real kinetic energy  $T[n]$ .

$$T_s[n] = \langle \Psi_s | \sum_i^N -\frac{\hbar^2}{2m_e} \nabla_i^2 | \Psi_s \rangle = \sum_i^N \langle \psi_i | -\frac{\hbar^2}{2m_e} \nabla^2 | \psi_i \rangle. \quad (5.8)$$

This results in an exact theory which is of one-particle form. To separate  $T_s[n]$  as the kinetic energy component,  $F[n]$  of equation (5.5) is rewritten to

$$F[n] = T_s[n] + J[n] + E_{xc}[n] \quad (5.9)$$

<sup>2</sup>Differences between the spin orbitals ( $s$  in  $\psi_i(\vec{r}, s)$ ) are omitted here, their effect will be discussed in section 5.5.3

<sup>3</sup>This is called  $N$ -representability.

with

$$J[n] = \frac{1}{8\pi\epsilon_0} \iint \frac{n(\vec{r})n(\vec{r}')}{|\vec{r} - \vec{r}'|} d^3r d^3r' \quad (5.10)$$

being the classical part of the interaction between the electrons<sup>4</sup>. All unknown parts are collected in

$$E_{xc}[n] = T[n] - T_s[n] + V_{ee} - J[n]. \quad (5.11)$$

$E_{xc}[n]$  is called the *exchange-correlation energy*. It contains the difference between  $T$  and  $T_s$ , and the non-classical part of the electron-electron interaction  $V_{ee}$ . This non-classical part of  $V_{ee}$  results in a so-called exchange-correlation hole caused by the Pauli-repulsion of the electrons. The one-particle states of the non-interacting reference system solve  $N$  equations

$$\left(-\frac{\hbar^2}{2m_e}\nabla^2 + v_{eff}(\vec{r})\right)\psi_i = \varepsilon_i\psi_i \quad (5.12)$$

with

$$n(\vec{r}) = \sum_i^N \sum_s |\psi_i(\vec{r}, s)|^2 \quad (5.13)$$

and the Kohn-Sham effective potential

$$\begin{aligned} v_{eff}(\vec{r}) &= v_{ext}(\vec{r}) + \frac{\delta J[n]}{\delta n(\vec{r})} + \frac{\delta E_{xc}[n]}{\delta n(\vec{r})} \\ &= v_{ext}(\vec{r}) + \frac{1}{4\pi\epsilon_0} \int \frac{n(\vec{r}')}{|\vec{r} - \vec{r}'|} d^3r' + v_{xc}(\vec{r}) \end{aligned} \quad (5.14)$$

with the exchange-correlation potential

$$v_{xc}(\vec{r}) = \frac{\delta E_{xc}[n]}{\delta n(\vec{r})}. \quad (5.15)$$

These equations (5.12)–(5.15) are the so-called Kohn-Sham equations. They have to be solved by iterating until self-consistency is reached. Orthogonality between the  $\psi_i$  has to be ensured to keep equation (5.8) valid. The total energy functional is

$$E[n] = T_s[n] + J[n] + \int v_{ext}(\vec{r})n(\vec{r}) d^3r + E_{xc}[n]. \quad (5.16)$$

All terms of this equation but  $E_{xc}[n]$ , the exchange and correlation energy, can be calculated exactly. The Kohn-Sham Hamiltonian is the sum of the one-particle Hamiltonians in equation (5.12) and acting on the Kohn-Sham wave function

---

<sup>4</sup>Hartree term

(5.7). The total energy is the sum of the one-particle energies of the occupied states. Detailed derivation and discussion of the Kohn-Sham equations can be found elsewhere [114, 116].

Given the auxiliary nature of the one-particle Kohn-Sham orbitals—just  $N$  orbitals the sum of squares of which add up to the true total electron density—there is no simple physical meaning for them and the one-particle energies.

### 5.1.3 Local density approximation (LDA)

While density functional theory in the Kohn-Sham formulation is able to provide the major part of the total energy exactly, an explicit approximation for  $E_{xc}$  is required for equation 5.16. The search more accurate functionals is a subject of ongoing research.

The functional for exchange and correlation energy is replaced by a local function. The simplest approximation for this function is to use the exchange and correlation energy of a uniform electron gas with the same density as the real system at that point. This is called the local density approximation (LDA):

$$E_{xc}^{LDA}[n] = \int n(\vec{r}) \epsilon_{xc}(n) d^3r \quad (5.17)$$

where  $\epsilon_{xc}(n)$  is the exchange and correlation energy per particle of a uniform electron gas of density  $n$ . It can be divided into exchange and correlation contributions:

$$\epsilon_{xc}(n) = \epsilon_x(n) + \epsilon_c(n). \quad (5.18)$$

The exchange part is, by definition, the part treated correctly in the Hartree-Fock method<sup>5</sup> representing the Pauli-repulsion of the electrons. For a uniform electron gas it can be calculated exactly from the Dirac exchange-energy functional:

$$\epsilon_x^{LDA}(n) = -\frac{e^2}{4\pi\epsilon_0} \frac{3}{4} \left(\frac{3}{\pi}\right)^{\frac{1}{3}} n(\vec{r})^{\frac{1}{3}}. \quad (5.19)$$

The correlation part  $\epsilon_c(n)$  cannot be calculated analytically even though its high-density limit and its low-density limit are known. These and other boundary conditions as well as the results of quantum Monte Carlo calculations have been used to parameterize  $\epsilon_c$ . Its effect on the total result is less pronounced because correlation energies are an order of magnitude smaller than exchange energies [114]. Throughout the simulations in this work, the parameterization of Perdew and Wang [117] will be used for the local correlation energy.

Even though the generalization from a uniform electron gas to the strongly varying density in atoms and molecules seems to be crude, LDA enables astonishingly

---

<sup>5</sup>for the free electron gas

accurate calculations and increased the acceptance of DFT in its first years. However, it has some shortcomings, for example it is known to overestimate binding energies [118–120].

#### 5.1.4 PBE gradient approximation

An improvement of the exchange-correlation functional LDA is to not only include the local density into the functional, but also its gradient:

$$E_{xc}^{GGA}[n] = \int n(\vec{r}) \epsilon_{xc}^{GGA}(n, |\nabla n|) d^3r. \quad (5.20)$$

This method is called generalized gradient approximation (GGA). It employs non-local functionals. Most of these functionals are based on LDA and add gradient corrections.

While some of the available parameterizations of GGA are fitted to experimental values, there are also a number of physical boundary conditions [114, 116] available for finding a good functional. In 1996 Perdew Burke and Ernzerhof [121] presented a parameter-free GGA form (PBE) which fulfills some of these boundary conditions. The PBE correlation and exchange energies are<sup>6</sup>:

$$\epsilon_c^{PBE} = \epsilon_c^{LDA} - \gamma \phi^3 \ln \left[ 1 + \frac{1}{\chi s^2 / \phi^2 + (\chi s^2 / \phi^2)^2} \right] \quad (5.21)$$

$$\gamma = \frac{1 - \ln 2}{\pi^2} \approx 0.031091,$$

$$\phi(\zeta) = \frac{1}{2} \left[ (1 + \zeta)^{2/3} + (1 - \zeta)^{2/3} \right], \quad \zeta = \frac{n_{\uparrow} - n_{\downarrow}}{n}$$

$$\chi \approx 0.72161, \quad s = \frac{|\nabla n|}{2n(3\pi^2 n)^{1/3}}$$

$$\epsilon_x^{PBE} = \epsilon_x^{LDA} \cdot F_x(s) \quad (5.22)$$

$$F_x(s) = 1 + \kappa - \frac{\kappa}{1 - \mu s^2 / \kappa} \quad (5.23)$$

$$\kappa = 0.804, \quad \mu \approx 0.21951$$

As this functional is not fitted to experimental data it can be expected to be valid for a wide range of systems. Experimental data are reproduced significantly better by PBE than by LDA [121].

While the most accurate functionals are constructed so that they match a number of physical boundary conditions, some arbitrariness remains in them. Therefore the accuracy of a functional depends on the system it is applied on. However, optimizing a functional for the system of interest may cause problems, as there is

<sup>6</sup>In SI units,  $\gamma$  should be multiplied with  $\frac{\hbar^2}{m_e \sigma_B^2} = (m_e^3 (4\pi\epsilon_0)^2 \hbar^2 e^4)^{-1}$



no guarantee that this optimized version performs well for a different, even similar system. I used PBE gradient correction for all calculations in this work, except expecially noted, as I wanted to choose a functional of general applicability that is widely distributed an known to give very accurate results for a wide range of systems.

## 5.2 The projector augmented wave method

One of the topics I was asked repeatedly during discussions with students and fellow scientists, was the electronic structure method and the basis set which I used: the Projector Augmented Wave (PAW) method. Therefore, Peter E. Blöchl, Clemens J. Först and myself wrote a paper [122] and later on a chapter in a book [123] explaining it. I also decided to include a section about PAW into this thesis.

I will separate this section into two parts, the first one giving a general overview understandable to the non-expert, and the second one providing more detailed and mathematical insight. The second part is mostly taken from the paper mentioned above [122].

### 5.2.1 PAW – The overview

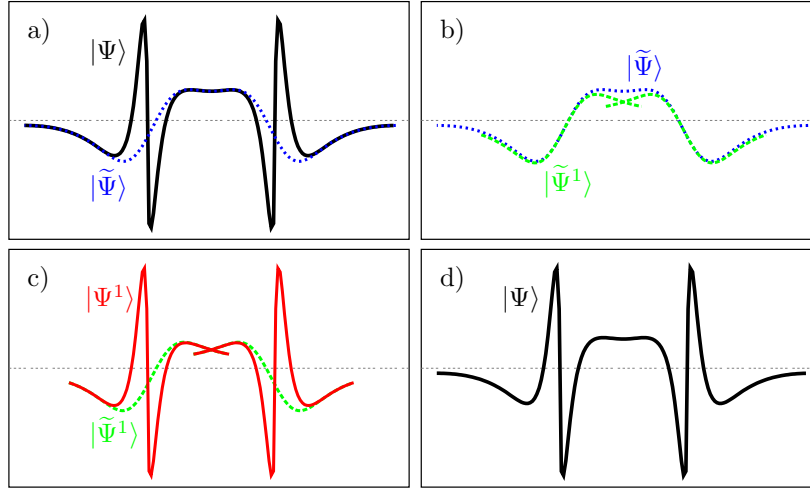
Every basis set used for electronic structure calculations has to deal with two major numerical problems: (1) In the atomic region near the nucleus, the kinetic energy of the electrons is large, resulting in rapid oscillations of the wave function, which require fine grids for an accurate numerical representation. On the other hand, the large kinetic energy makes the wave function stiff, so that its shape is hardly influenced by changes in the chemical environment. Therefore, a small basis set sufficiently represents the wave function in the atomic region. (2) In the interstitial region between the atoms the situation is opposite. The kinetic energy is small, therefore the wave function is smooth. However, the wave function is flexible and responds strongly to the environment as this is the region of chemical bonds. This requires large and nearly complete basis sets.

The interstitial region is best described with plane waves able to describe any smooth wavefunction and effectively treatable on computers. The atomic region requires some augmentation of the plane waves, describing the spikes and nodes.

Fig. 5.1a on the facing page shows the real, physical wave function  $|\Psi\rangle$  of a  $\sigma$ -bonding orbital of a chlorine dimer. It also shows a smooth dotted line for the auxiliary wave function  $|\tilde{\Psi}\rangle$  used in the PAW method.  $|\tilde{\Psi}\rangle$  exactly matches the real wavefunction in the interstitial region but smoothes it within the atoms. A tilde ( $\sim$ ) denotes auxiliary functions, indicating their smooth behavior<sup>7</sup>.

To correct the error of the auxiliary wave function within the atomic region its one-center expansion  $|\tilde{\Psi}^1\rangle$  within each atom is calculated. This is illustrated in Fig. 5.1b. A mathematical algorithm, the heart of the PAW method, allows to obtain the one-center expansion of the real wave function  $|\Psi^1\rangle$  from  $|\tilde{\Psi}^1\rangle$  as depicted in Fig. 5.1c.

<sup>7</sup>The literature [122, 124, 125] sometimes refers to the auxiliary wave function as pseudo wave function while referring to the real wave function as all-electron wave function.



**Figure 5.1** The PAW wave functions: a) The real and auxiliary wave function. b) The auxiliary wave function and its one-center expansion. c) The two one-center expansions and d) the real wave function.

Adding the difference of these one-center expansions to the auxiliary wave function leads to the real wave function, shown in Fig. 5.1d. All physical properties are calculated from this real wave function.

$$|\Psi\rangle = |\tilde{\Psi}\rangle + (|\Psi^1\rangle - |\tilde{\Psi}^1\rangle) \quad (5.24)$$

This formalism is separately applied to each atom in each one-particle state<sup>8</sup> (each orbital) and to the electron density. The density of the core electrons is added to the valence density.

The PAW method has some similarities to the pseudopotential method, as the variational parameters are plane wave coefficients. However, it reconstructs the real wave function and calculates all physical properties out of that. It can be viewed as a pseudopotential method with pseudopotentials instantaneously adopting to the chemical environment of the atom.

### 5.2.2 PAW – The mathematical basis

#### Transformation theory

The root of the PAW method is a transformation, that maps the true wave functions with their complete nodal structure onto auxiliary wave functions, that are

<sup>8</sup>Also to each spin direction or component and each  $k$ -point.

numerically convenient. We aim for smooth auxiliary wave functions, which have a rapidly convergent plane wave expansion. With such a transformation I can expand the auxiliary wave functions into a convenient basis set, and evaluate all physical properties after reconstructing the related physical (true) wave functions.

Let us denote the physical one-particle wave function as  $|\Psi\rangle$  and the auxiliary wave function as  $|\tilde{\Psi}\rangle$ . Note that the tilde refers to the representation of smooth auxiliary wave functions. These wave functions have to be constructed separately for each one-particle state,  $k$ -point and spin direction or component. The transformation from the auxiliary to the physical wave functions is  $\mathcal{T}$ .

$$|\Psi\rangle = \mathcal{T}|\tilde{\Psi}\rangle \quad (5.25)$$

Thus the Kohn-Sham equation

$$\mathcal{H}|\Psi\rangle = |\Psi\rangle\epsilon \quad (5.26)$$

takes the form

$$\mathcal{T}^\dagger \mathcal{H} \mathcal{T} |\tilde{\Psi}\rangle = \mathcal{T}^\dagger \mathcal{T} |\tilde{\Psi}\rangle \epsilon. \quad (5.27)$$

Again we obtain a Schrödinger-like equation, but now the Hamilton operator has a different form,  $\mathcal{T}^\dagger \mathcal{H} \mathcal{T}$ , an overlap operator  $\mathcal{T}^\dagger \mathcal{T}$  occurs. The resulting auxiliary wave functions are smooth.

When we evaluate physical quantities we need to evaluate expectation values of an operator  $\mathcal{A}$ , which can be expressed in terms of either the true or the auxiliary wave functions.

$$\langle \mathcal{A} \rangle = \sum_n f_n \langle \Psi_n | \mathcal{A} | \Psi_n \rangle = \sum_n f_n \langle \tilde{\Psi}_n | \mathcal{T}^\dagger \mathcal{A} \mathcal{T} | \tilde{\Psi}_n \rangle \quad (5.28)$$

Here,  $n$  numbers the band,  $k$ -point and spin. In the representation of auxiliary wave functions we need to use transformed operators  $\tilde{\mathcal{A}} = \mathcal{T}^\dagger \mathcal{A} \mathcal{T}$ . As it is, this equation only holds for the valence electrons. The core electrons are treated differently as will be shown below.

The transformation takes us conceptionally from the world of pseudopotentials to that of augmented wave methods, which deal with the full wave functions. We will see that our auxiliary wave functions, which are simply the plane wave parts of the full wave functions, translate into the wave functions of the pseudopotential approach. In the PAW method the auxiliary wave functions are used to construct the true wave functions and the total energy functional is evaluated from the latter. Thus it provides the missing link between augmented wave methods and the pseudopotential method, which can be derived as a well-defined approximation of the PAW method.

In the original paper [124], the auxiliary wave functions have been termed pseudo wave functions and the true wave functions have been termed all-electron wave functions, in order to make the connection more evident. I avoid this notation here, because it resulted in confusion in cases, where the correspondence is not clear cut.

### Transformation operator

Sofar, I have described how to determine the auxiliary wave functions of the ground state and how to obtain physical information from them. What is missing, is a definition of the transformation operator  $\mathcal{T}$ .

The operator  $\mathcal{T}$  has to modify the smooth auxiliary wave function in each atomic region, so that the resulting wave function has the correct nodal structure. Therefore, it makes sense to write the transformation as identity plus a sum of atomic contributions  $\mathcal{S}_R$

$$\mathcal{T} = 1 + \sum_R \mathcal{S}_R. \quad (5.29)$$

The index  $R$  is a label for an atomic site. For every atom,  $\mathcal{S}_R$  adds the difference between the true and the auxiliary wave function.

The local terms  $\mathcal{S}_R$  are defined in terms of solutions  $|\phi_i\rangle$  of the Schrödinger equation for the isolated atoms. This set of partial waves  $|\phi_i\rangle$  will serve as a basis set so that, near the nucleus, all relevant valence wave functions can be expressed as superposition of the partial waves with yet unknown coefficients.

$$\Psi(\vec{r}) = \sum_{i \in R} \phi_i(\vec{r}) c_i \quad \text{for} \quad |\vec{r} - \vec{R}_R| < r_{c,R} \quad (5.30)$$

The index  $i$  refers to a site index  $R$ , the angular momentum indices  $(\ell, m)$  and an additional index that differentiates partial waves with same angular momentum quantum numbers on the same site. The partial waves that belonging to site  $R$  are indicated by  $i \in R$ .  $\vec{R}_R$  is the position of the nucleus of site  $R$ .

Note that the partial waves are not necessarily bound states and are therefore not normalizable, unless we truncate them beyond a certain radius  $r_{c,R}$ . The PAW method is formulated such that the final results do not depend on the location where the partial waves are truncated, as long as this is not done too close to the nucleus.

Since the core wave functions do not spread out into the neighboring atoms, I will treat them differently. Currently I use the frozen-core approximation so that density and energy of the core electrons are identical to those of the corresponding isolated atoms. The transformation  $\mathcal{T}$  shall produce only wave functions orthogonal to the core electrons, while the core electrons are treated separately.

Therefore, the set of atomic partial waves  $|\phi_i\rangle$  includes only valence states that are orthogonal to the core wave functions of the atom.

For each of the partial waves we choose an auxiliary partial wave  $|\tilde{\phi}_i\rangle$ . The identity

$$\begin{aligned} |\phi_i\rangle &= (1 + \mathcal{S}_R)|\tilde{\phi}_i\rangle \quad \text{for } i \in R \\ \mathcal{S}_R|\tilde{\phi}_i\rangle &= |\phi_i\rangle - |\tilde{\phi}_i\rangle \end{aligned} \quad (5.31)$$

defines the local contribution  $\mathcal{S}_R$  to the transformation operator. Since  $1 + \mathcal{S}_R$  shall change the wave function only locally, we require that the partial waves  $|\phi_i\rangle$  and their auxiliary counter parts  $|\tilde{\phi}_i\rangle$  are pairwise identical beyond a certain radius  $r_c$ .

$$\phi_i(\vec{r}) = \tilde{\phi}_i(\vec{r}) \quad \text{for } i \in R \quad \text{and} \quad |\vec{r} - \vec{R}_R| > r_{c,R} \quad (5.32)$$

In order to be able to apply the transformation operator to an arbitrary auxiliary wave function, we need to be able to expand the auxiliary wave function locally into the auxiliary partial waves.

$$\tilde{\Psi}(\vec{r}) = \sum_{i \in R} \tilde{\phi}_i(\vec{r}) \langle \tilde{p}_i | \tilde{\Psi} \rangle \quad \text{for } |\vec{r} - \vec{R}_R| < r_{c,R} \quad (5.33)$$

which defines the projector functions  $|\tilde{p}_i\rangle$ . The projector functions probe the local character of the auxiliary wave function in the atomic region. Examples of projector functions are shown in Fig. 5.2 on the next page. From equation (5.33) we can derive  $\sum_i |\tilde{\phi}_i\rangle \langle \tilde{p}_i| = 1$ , which is valid within  $r_c$ . It can be shown by insertion, that the identity equation (5.33) holds for any auxiliary wave function  $|\tilde{\Psi}\rangle$  that can be expanded locally into auxiliary partial waves  $|\tilde{\phi}_i\rangle$ , if

$$\langle \tilde{p}_i | \tilde{\phi}_j \rangle = \delta_{i,j} \quad \text{for } i, j \in R \quad (5.34)$$

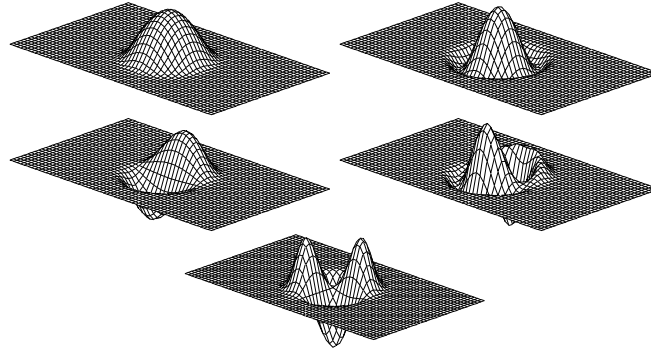
Note that neither the projector functions nor the partial waves need to be orthogonal among themselves.

By combining the equations (5.31) and (5.33), we can apply  $\mathcal{S}_R$  to any auxiliary wave function.

$$\mathcal{S}_R|\tilde{\Psi}\rangle = \sum_{i \in R} \mathcal{S}_R|\tilde{\phi}_i\rangle \langle \tilde{p}_i | \tilde{\Psi} \rangle = \sum_{i \in R} (|\phi_i\rangle - |\tilde{\phi}_i\rangle) \langle \tilde{p}_i | \tilde{\Psi} \rangle \quad (5.35)$$

Hence the transformation operator is

$$\mathcal{T} = 1 + \sum_i (|\phi_i\rangle - |\tilde{\phi}_i\rangle) \langle \tilde{p}_i | \quad (5.36)$$



**Figure 5.2** Top: projector functions of the chlorine atom for two s-type partial waves, middle: p-type, bottom: d-type.

where the sum runs over all partial waves of all atoms. The true wave function can be expressed as

$$|\Psi\rangle = |\tilde{\Psi}\rangle + \sum_i (|\phi_i\rangle - |\tilde{\phi}_i\rangle) \langle \tilde{p}_i | \tilde{\Psi} \rangle = |\tilde{\Psi}\rangle + \sum_R (|\Psi_R^1\rangle - |\tilde{\Psi}_R^1\rangle) \quad (5.37)$$

with

$$|\Psi_R^1\rangle = \sum_{i \in R} |\phi_i\rangle \langle \tilde{p}_i | \tilde{\Psi} \rangle \quad (5.38)$$

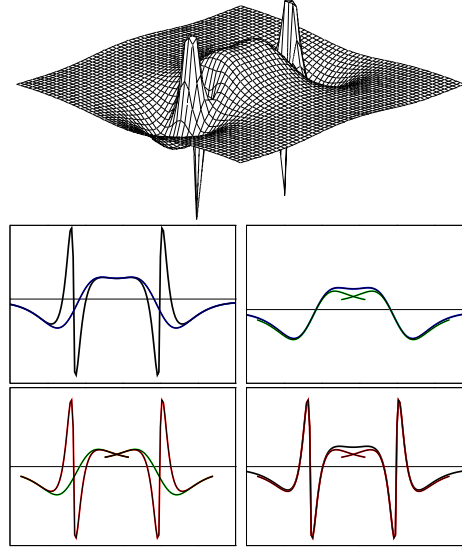
$$|\tilde{\Psi}_R^1\rangle = \sum_{i \in R} |\tilde{\phi}_i\rangle \langle \tilde{p}_i | \tilde{\Psi} \rangle \quad (5.39)$$

In Fig. 5.3 on the following page the decomposition of equation (5.37) is shown for the example of the bonding p- $\sigma$  state of a chlorine molecule.

To understand the expression for the true wave function (5.37) let us concentrate on different regions in space. (1) Far from the atoms, the partial waves are, according to equation (5.32), pairwise identical so that the auxiliary wave function is identical to the true wave function  $\Psi(\vec{r}) = \tilde{\Psi}(\vec{r})$ . (2) Close to an atom, however, the true wave function  $\Psi(\vec{r}) = \Psi_R^1(\vec{r})$  is built up from partial waves that contain the proper nodal structure, because the auxiliary wave function and its partial wave expansion are equal, according to equation (5.33).

In practice the partial wave expansions are truncated. Therefore, the identity of equation (5.33) does not hold strictly. As a result the plane waves also contribute to the true wave function inside the atomic region. This has the advantage that the missing terms in a truncated partial wave expansion are partly accounted for by plane waves, which explains the rapid convergence of the partial wave expansions.

Frequently, the question comes up, whether the transformation equation (5.36) of the auxiliary wave functions indeed provides the true wave function. The transformation should be considered merely as a change of representation analogous to



**Figure 5.3** Bonding p- $\sigma$  orbital of the  $\text{Cl}_2$  molecule and its decomposition of the wave function into auxiliary wave function and the two one-center expansions. Top-left: true and auxiliary wave function; top-right: auxiliary wave function and its partial wave expansion; bottom-left: the two partial wave expansions; bottom-right: true wave function and its partial wave expansion.

a coordinate transform. If the total energy functional is transformed consistently, its minimum will yield an auxiliary wave function that produces a correct wave function  $|\Psi\rangle$ .

### Expectation values

Expectation values can be obtained either from the reconstructed true wave functions or directly from the auxiliary wave functions

$$\begin{aligned}
 \langle \mathcal{A} \rangle &= \sum_n f_n \langle \Psi_n | \mathcal{A} | \Psi_n \rangle + \sum_{n=1}^{N_c} \langle \phi_n^c | \mathcal{A} | \phi_n^c \rangle \\
 &= \sum_n f_n \langle \tilde{\Psi}_n | \mathcal{T}^\dagger \mathcal{A} \mathcal{T} | \tilde{\Psi}_n \rangle + \sum_{n=1}^{N_c} \langle \phi_n^c | \mathcal{A} | \phi_n^c \rangle
 \end{aligned} \tag{5.40}$$

where  $f_n$  are the occupations of the valence states and  $N_c$  is the number of core states. The first sum runs over the valence states, and second over the core states  $|\phi_n^c\rangle$ .

Now we can decompose the matrix elements into their individual contributions



according to (5.37).

$$\begin{aligned}
\langle \Psi | \mathcal{A} | \Psi \rangle &= \left\langle \tilde{\Psi} + \sum_R (\Psi_R^1 - \tilde{\Psi}_R^1) \middle| \mathcal{A} \middle| \tilde{\Psi} + \sum_{R'} (\Psi_{R'}^1 - \tilde{\Psi}_{R'}^1) \right\rangle \\
&= \underbrace{\langle \tilde{\Psi} | \mathcal{A} | \tilde{\Psi} \rangle + \sum_R \left( \langle \Psi_R^1 | \mathcal{A} | \Psi_R^1 \rangle - \langle \tilde{\Psi}_R^1 | \mathcal{A} | \tilde{\Psi}_R^1 \rangle \right)}_{\text{part 1}} \\
&\quad + \underbrace{\sum_R \left( \langle \Psi_R^1 - \tilde{\Psi}_R^1 | \mathcal{A} | \tilde{\Psi} - \tilde{\Psi}_R^1 \rangle + \langle \tilde{\Psi} - \tilde{\Psi}_R^1 | \mathcal{A} | \Psi_R^1 - \tilde{\Psi}_R^1 \rangle \right)}_{\text{part 2}} \\
&\quad + \underbrace{\sum_{R \neq R'} \langle \Psi_R^1 - \tilde{\Psi}_R^1 | \mathcal{A} | \Psi_{R'}^1 - \tilde{\Psi}_{R'}^1 \rangle}_{\text{part 3}} \tag{5.41}
\end{aligned}$$

Only the first part of equation (5.41), is evaluated explicitly, while the second and third parts of equation (5.41) are neglected, because they vanish for sufficiently local operators as long as the partial wave expansion is converged: the function  $\Psi_R^1 - \tilde{\Psi}_R^1$  vanishes per construction beyond some augmentation region, because the partial waves are pairwise identical beyond that region. The function  $\tilde{\Psi} - \tilde{\Psi}_R^1$  vanishes inside the augmentation region, if the partial wave expansion is sufficiently converged. In no region of space both functions  $\Psi_R^1 - \tilde{\Psi}_R^1$  and  $\tilde{\Psi} - \tilde{\Psi}_R^1$  are simultaneously non-zero. Similarly the functions  $\Psi_R^1 - \tilde{\Psi}_R^1$  from different sites are never non-zero in the same region in space. Hence, the second and third parts of equation (5.41) vanish for operators such as the kinetic energy  $\frac{-\hbar^2}{2m_e} \nabla^2$  and the real space projection operator  $|r\rangle\langle r|$ , which produces the electron density. For truly non-local operators the second and third parts of equation (5.41) would have to be considered explicitly.

The expression (5.41) for the expectation value can therefore be written as

$$\begin{aligned}
\langle \mathcal{A} \rangle &= \sum_n f_n \left( \langle \tilde{\Psi}_n | \mathcal{A} | \tilde{\Psi}_n \rangle + \langle \Psi_n^1 | \mathcal{A} | \Psi_n^1 \rangle - \langle \tilde{\Psi}_n^1 | \mathcal{A} | \tilde{\Psi}_n^1 \rangle \right) + \sum_{n=1}^{N_c} \langle \phi_n^c | \mathcal{A} | \phi_n^c \rangle \\
&= \sum_n f_n \langle \tilde{\Psi}_n | \mathcal{A} | \tilde{\Psi}_n \rangle + \sum_{n=1}^{N_c} \langle \tilde{\phi}_n^c | \mathcal{A} | \tilde{\phi}_n^c \rangle \\
&\quad + \sum_R \left( \sum_{i,j \in R} D_{i,j} \langle \phi_j | \mathcal{A} | \phi_i \rangle + \sum_{n \in R}^{N_{c,R}} \langle \phi_n^c | \mathcal{A} | \phi_n^c \rangle \right) \\
&\quad - \sum_R \left( \sum_{i,j \in R} D_{i,j} \langle \tilde{\phi}_j | \mathcal{A} | \tilde{\phi}_i \rangle + \sum_{n \in R}^{N_{c,R}} \langle \tilde{\phi}_n^c | \mathcal{A} | \tilde{\phi}_n^c \rangle \right) \tag{5.42}
\end{aligned}$$

where  $D_{i,j}$  is the one-center density matrix defined as

$$D_{i,j} = \sum_n f_n \langle \tilde{\Psi}_n | \tilde{p}_j \rangle \langle \tilde{p}_i | \tilde{\Psi}_n \rangle = \sum_n \langle \tilde{p}_i | \tilde{\Psi}_n \rangle f_n \langle \tilde{\Psi}_n | \tilde{p}_j \rangle \quad (5.43)$$

The auxiliary core states,  $|\tilde{\phi}_n^c\rangle$  allow to incorporate the tails of the core wave function into the plane wave part, and therefore assure, that the integrations of partial wave contributions cancel strictly beyond  $r_c$ . They are identical to the true core states in the tails, but are a smooth continuation inside the atomic sphere. It is not required that the auxiliary wave functions are normalized.

For example, the electron density is given by

$$n(\vec{r}) = \tilde{n}(\vec{r}) + \sum_R (n_R^1(\vec{r}) - \tilde{n}_R^1(\vec{r})) \quad (5.44)$$

$$\begin{aligned} \tilde{n}(\vec{r}) &= \sum_n f_n \tilde{\Psi}_n^*(\vec{r}) \tilde{\Psi}_n(\vec{r}) + \tilde{n}_c \\ n_R^1(\vec{r}) &= \sum_{i,j \in R} D_{i,j} \phi_j^*(\vec{r}) \phi_i(\vec{r}) + n_{c,R} \\ \tilde{n}_R^1(\vec{r}) &= \sum_{i,j \in R} D_{i,j} \tilde{\phi}_j^*(\vec{r}) \tilde{\phi}_i(\vec{r}) + \tilde{n}_{c,R} \end{aligned} \quad (5.45)$$

where  $n_{c,R}$  is the core density of the corresponding atom and  $\tilde{n}_{c,R}$  is the auxiliary core density that is identical to  $n_{c,R}$  outside the atomic region and a smooth continuation inside.

Before I continue, let us discuss a special point: the matrix element of a general operator with the auxiliary wave functions may be slowly converging with the plane wave expansion, because the operator  $\mathcal{A}$  may not be well behaved. An example for such an operator is the singular electrostatic potential of a nucleus. This problem can be alleviated by adding an intelligent zero: if an operator  $\mathcal{B}$  is purely localized within an atomic region, we can use the identity between the auxiliary wave function and its own partial wave expansion

$$0 = \langle \tilde{\Psi}_n | \mathcal{B} | \tilde{\Psi}_n \rangle - \langle \tilde{\Psi}_n^1 | \mathcal{B} | \tilde{\Psi}_n^1 \rangle \quad (5.46)$$

Now we choose an operator  $\mathcal{B}$  so that it cancels the problematic behavior of the operator  $\mathcal{A}$ , but is localized in a single atomic region. By adding  $\mathcal{B}$  to the plane wave part and the matrix elements with its one-center expansions, the plane wave convergence can be improved without affecting the converged result.

### Total energy

Like wave functions and expectation values also the total energy can be divided into three parts.

$$E([\tilde{\Psi}_n], R_i) = \tilde{E} + \sum_R (E_R^1 - \tilde{E}_R^1) \quad (5.47)$$

The plane-wave part  $\tilde{E}$  involves only smooth functions and is evaluated on equispaced grids in real and reciprocal space. This part is computationally most demanding, and is similar to the expressions in the pseudopotential approach.

$$\begin{aligned} \tilde{E} &= \sum_n \langle \tilde{\Psi}_n | \frac{\hbar^2}{2m_e} \nabla^2 | \tilde{\Psi}_n \rangle \\ &+ \frac{e^2}{8\pi\epsilon_0} \int d^3r \int d^3r' \frac{(\tilde{n}(\vec{r}) + \tilde{Z}(\vec{r}))(\tilde{n}(\vec{r}') + \tilde{Z}(\vec{r}'))}{|\vec{r} - \vec{r}'|} \\ &+ \int d^3r \tilde{n}(\vec{r}) \epsilon_{xc}(\vec{r}, [\tilde{n}]) + \int d^3r \tilde{v}(\vec{r}) \tilde{n}(\vec{r}), \end{aligned} \quad (5.48)$$

where  $\tilde{Z}(\vec{r})$  is an angular dependent core-like density that will be described in detail below. The remaining parts can be evaluated on radial grids in a spherical harmonics expansion. The nodal structure of the wave functions can be properly described on a logarithmic radial grid that becomes very fine near the nucleus,

$$\begin{aligned} E_R^1 &= \sum_{i,j \in R} D_{i,j} \langle \phi_j | \frac{\hbar^2}{2m_e} \nabla^2 | \phi_i \rangle + \sum_{n \in R}^{N_{c,R}} \langle \phi_n^c | \frac{\hbar^2}{2m_e} \nabla^2 | \phi_n^c \rangle \\ &+ \frac{e^2}{8\pi\epsilon_0} \int d^3r \int d^3r' \frac{(n^1(\vec{r}) + Z(\vec{r}))(n^1(\vec{r}') + Z(\vec{r}'))}{|\vec{r} - \vec{r}'|} \\ &+ \int d^3r n^1(\vec{r}) \epsilon_{xc}(\vec{r}, [n^1]) \end{aligned} \quad (5.49)$$

$$\begin{aligned} \tilde{E}_R^1 &= \sum_{i,j \in R} D_{i,j} \langle \tilde{\phi}_j | \frac{\hbar^2}{2m_e} \nabla^2 | \tilde{\phi}_i \rangle \\ &+ \frac{e^2}{8\pi\epsilon_0} \int d^3r \int d^3r' \frac{(\tilde{n}^1(\vec{r}) + \tilde{Z}(\vec{r}))(\tilde{n}^1(\vec{r}') + \tilde{Z}(\vec{r}'))}{|\vec{r} - \vec{r}'|} \\ &+ \int d^3r \tilde{n}^1(\vec{r}) \epsilon_{xc}(\vec{r}, [\tilde{n}^1]) + \int d^3r \tilde{v}(\vec{r}) \tilde{n}^1(\vec{r}) \end{aligned} \quad (5.50)$$

The nuclear charge density  $-eZ(\vec{r})$  is defined as a sum of  $\delta$ -functions on the nuclear sites,  $Z(\vec{r}) = -\sum_R Z_R \delta(\vec{r} - \vec{R})$ , with the atomic numbers  $Z_R$ . Note that the self energy of a point charge is infinite and must be subtracted out.

The compensation density  $\tilde{Z}(\vec{r}) = \sum_R \tilde{Z}_R(\vec{r})$  is given as a sum of angular momentum dependent Gauss functions, which have an analytical Fourier transform. A similar term occurs also in the pseudopotential approach. In contrast to the norm-conserving pseudopotential approach however, the compensation charge is non-spherical and it is constantly adapting to the instantaneous environment. It is constructed such that the augmentation charge densities

$$n_R^1(\vec{r}) + Z_R(\vec{r}) - \tilde{n}_R^1(\vec{r}) - \tilde{Z}_R(\vec{r}) \quad (5.51)$$

have vanishing electrostatic multi-pole moments for each atomic site. As a result the sum of all one-center contributions from one atom does not produce an electrostatic potential outside their own atomic region. This is the reason why the electrostatic interaction of the one-center parts between different sites vanish.

The compensation charge density as given here is still localized within the atomic regions, but a technique similar to an Ewald summation allows to replace it by a very extended charge density. Thus we can achieve, that all functions in  $\tilde{E}$  converge as fast as the auxiliary density itself.

The potential  $\bar{v}$ , which occurs in equations (5.48) and (5.50) enters the total energy in the form of a zero described in equation (5.46)

$$\sum_n f_n \langle \tilde{\Psi}_n | \left( \bar{v} - \sum_{i,j} |\tilde{p}_i\rangle \langle \tilde{\phi}_i | \bar{v} | \tilde{\phi}_j\rangle \langle \tilde{p}_j| \right) | \tilde{\Psi}_n \rangle \quad (5.52)$$

The potential  $\bar{v} = \sum_R \bar{v}_R$ , which occurs in equations (5.48) and (5.50) enters the total energy in the form of “intelligent zeros” described in equation (5.46)

$$\begin{aligned} 0 &= \sum_n f_n \left( \langle \tilde{\Psi}_n | \bar{v}_R | \tilde{\Psi}_n \rangle - \langle \tilde{\Psi}_n^1 | \bar{v}_R | \tilde{\Psi}_n^1 \rangle \right) \\ &= \sum_n f_n \langle \tilde{\Psi}_n | \bar{v}_R | \tilde{\Psi}_n \rangle - \sum_{i,j \in R} D_{i,j} \langle \tilde{\phi}_i | \bar{v}_R | \tilde{\phi}_j \rangle \end{aligned} \quad (5.53)$$

The main reason for introducing this potential is that the self-consistent potential resulting from the plane wave part is not necessarily optimally smooth. The potential  $\bar{v}$  allows to influence the plane wave convergence beneficially, without changing the converged result.  $\bar{v}$  must be localized within the augmentation region, where equation (5.33) holds.

### Approximations

Once the total energy functional provided in the previous section has been defined, everything else follows: forces are partial derivatives with respect to atomic positions. The potential is the derivative of the potential energy with respect to

the density, and the Hamiltonian follows from derivatives with respect to wave functions. The fictitious Lagrangian approach of Car and Parrinello, discussed in section 5.3.2 on page 53, does not allow any freedom in the way these derivatives are obtained. Anything else than analytic derivatives will violate energy conservation in a dynamical simulation. Since the expressions are straightforward, even though rather involved, I will not discuss them here.

All approximations are incorporated already in the total energy functional of the PAW method. What are those approximations?

- The frozen core approximation. In principle this approximation can be overcome.
- The plane wave expansion for the auxiliary wave functions must be complete. This is easily controlled by increasing the plane wave cutoff defined as  $E_{PW} = \frac{1}{2}\hbar^2 G_{max}^2$ . Typically we use a plane wave cutoff of 30 Ry.
- The partial wave expansions must be converged. Typically we use one or two partial waves per angular momentum  $(\ell, m)$  and site. It should be noted that the partial wave expansion is not variational, because it changes the total energy functional and not only the basis set.

I do not discuss numerical approximations such as the choice of the radial grid, since those are easily controlled.

Earlier I mentioned that the pseudopotential approach can be derived as a well defined approximation from the PAW method: the augmentation part  $\Delta E = E^1 - \tilde{E}^1$  is a functional of the one-center density matrix  $D_{i,j}$  defined in equation (5.43). The pseudopotential approach can be recovered if we truncate a Taylor expansion of  $\Delta E$  about the atomic density matrix after the linear term. The term linear to  $D_{i,j}$  is the energy related to the non-local pseudopotential.

$$\begin{aligned} \Delta E(D_{i,j}) &= \Delta E(D_{i,j}^{at}) + \sum_{i,j} (D_{i,j} - D_{i,j}^{at}) \frac{\partial \Delta E}{\partial D_{i,j}} + O(D_{i,j} - D_{i,j}^{at})^2 \\ &= E_{self} + \sum_n f_n \langle \tilde{\Psi}_n | v_{nl} | \tilde{\Psi}_n \rangle + O(D_{i,j} - D_{i,j}^{at})^2 \end{aligned} \quad (5.54)$$

Thus we can look at the PAW method also as a pseudopotential method with a pseudopotential that adapts to the instantaneous electronic environment, because the explicit non-linear dependence of the total energy on the one-center density matrix is properly taken into account.

What are the main advantages of the PAW method compared to the pseudopotential approach?

Firstly all errors can be systematically controlled so that there are no transferability errors. As shown by Watson and Carter [126] and Kresse and Joubert

[125], most pseudopotentials fail for high spin atoms such as chromium. While it is probably true that pseudopotentials can be constructed that cope even with this situation, a failure can not be known beforehand, so that some empiricism remains in practice: a pseudopotential constructed from an isolated atom is not guaranteed to be accurate for a molecule. In contrast, the converged results of the PAW method do not depend on a reference system such as an isolated atom, because it uses the full density and potential.

The PAW method provides access to the full charge and spin density, which is relevant for hyperfine parameters. Hyperfine parameters are sensitive probes of the electron density near the nucleus. In many situations they are the only information available that allows to deduce atomic structure and chemical environment of an atom. There are reconstruction techniques for the pseudopotential approach [127], which however, are poor man's versions of the PAW method.

The plane wave convergence is more rapid than in norm-conserving pseudopotentials and should in principle be equivalent to that of ultra-soft pseudopotentials [128]. Compared to the ultra-soft pseudopotentials, however, the PAW method has the advantage that the total energy expression is less complex and therefore is expected to be more efficient.

### 5.3 Molecular dynamics

Molecular dynamics simulate the motion of atoms at a finite temperature. They may, however, also be used to find the ground state of a system. This can be achieved by applying friction to the atomic motion.

Knowledge of the forces acting on the atoms is required for molecular dynamics simulations. These forces may be calculated via  $F_i = \frac{\partial E}{\partial R_i}$  from the total energy of the system.  $E$  is obtained from the correct electronic structure based on DFT and the PAW method. Knowing these forces it is possible to calculate the minimum-energy structure, the ground state structure of the system.

#### 5.3.1 Discretization of the equations of motion

The motion of atoms obeys the laws of classical mechanics, expressed in Newton's equations of motion:

$$M_i \ddot{R}_i = F_i \quad (5.55)$$

with  $i$  numbering the  $3N$  degrees of freedom. Knowing the starting geometry and applying a friction to damp out the movement:

$$M_i \ddot{R}_i = F_i - M_i \alpha \dot{R}_i \quad (5.56)$$

the ground state structure can be calculated. To reach fast convergence, one starts with a high friction to damp out the fast vibration modes and lowers the friction continuously. This is necessary to also allow soft modes to move and finally damp out. In the current implementation of the program, the friction is set to a higher value whenever the total energy increases. This ensures that the atoms move towards the energy minimum.

Discretization of the equations of motion is necessary for simulations. Therefore, the trajectory is discretized in time steps  $\Delta t$ . The first and second derivatives with respect to time are replaced by the differential quotient. The scheme is called Verlet algorithm [129]:

$$\dot{R}_i(t) \approx \frac{R_i(t + \Delta t) - R_i(t - \Delta t)}{2 \Delta t} \quad (5.57)$$

$$\ddot{R}_i(t) \approx \frac{R_i(t + \Delta t) - 2R_i(t) + R_i(t - \Delta t)}{(\Delta t)^2} \quad (5.58)$$

While  $R_i(t)$  represents the current atomic positions,  $R_i(t - \Delta t)$  and  $R_i(t + \Delta t)$  represent the last and the following positions, respectively. This results in damped equations of motion as

$$R_i(t + \Delta t) = \frac{1}{1 + a} \left( 2R_i(t) - (1 - a)R_i(t - \Delta t) + \frac{F_i(\Delta t)^2}{M_i} \right) \quad (5.59)$$

with  $a = \alpha \Delta t / 2$ . The time step  $\Delta t$  has to be chosen according to the desired accuracy. The Verlet algorithm diverges for time steps larger than 30 % of the smallest vibration period of the system.

### 5.3.2 Ab-initio molecular dynamics

The direct application of equation (5.59) requires the determination of the electronic structure in each time step of nuclear motion. Car and Parrinello showed in 1985 [130] that simulation of simultaneous motion of electrons and nuclei is possible (*Car-Parrinello molecular dynamics*). In this approach, the electronic motion is not described by the time dependent Schrödinger-equation, but replaced by a fictitious motion keeping the electrons in the ground state:

$$m_\Psi |\ddot{\Psi}\rangle = -\frac{\partial E}{\partial \langle \Psi |} + \text{constraints} = -\mathcal{H}|\Psi\rangle + \text{constraints}. \quad (5.60)$$

A fictitious mass  $m_\Psi$  is assigned to the wave function. The forces acting on the wave functions are obtained from the energy functional.

Constraints ensure orthonormality of the wave functions. If the electronic ground state has been obtained, electronic and nuclear motion may be simulated with the same time step but different frictions.

In the following, I describe how the electronic and structural optimizations of this work have been performed. Random wave functions and a fixed atomic geometry have been used as starting conditions. First, the wave functions are optimized. The electronic ground state is reached with decreasing friction applied to the wave function dynamics.

With the correct electronic ground state, the atomic positions are optimized by dynamics under decreasing friction. Now, a rather high but also decreasing friction is applied to the electronic motion as the electrons are already in an optimized state. Whenever the energy increases, a high friction is applied to electrons ( $a = 0.03$ ) and nuclei ( $a = 0.003$ ). This ensures that the ground state is found.

Convergence of the structural relaxation is tested by simulating 200 time steps (50 fs) without any friction on the nuclei and a rather small one ( $a = 0.008$ ) on the wave functions. The system is considered to be converged if the kinetic temperature stays below 5 K and the total energy does not change significantly during this simulation.

Besides Car-Parrinello molecular dynamics, there are other methods for finding the electronic ground state [131] which I want to mention briefly. Apart from random starting conditions, the plane wave coefficients can be obtained from diagonalization of a plane-wave subset, or from diagonalization of a localized basis set and its expansion into plane waves [132]. The direct method to obtain an electronic ground state is diagonalization of the Hamilton matrix. In most cases, this is, however, computationally not feasible. The steepest descent scheme for finding the electronic ground state can be derived from the above described molecular dynamics by setting  $a = 1$ . It converges poorly. All-band conjugate gradient schemes [133] converge fast for non-metallic systems. They require no system-specific input parameters. In contrast to that, band-by-band conjugate gradient schemes [134, 135] also work for metallic systems. However, their stability depends on system-specific parameters obtained from preconditioning. Conjugate gradient schemes are also used for structural optimization.

## 5.4 Reaction rates

While some chemical reactions occur spontaneously, in most cases the educt and the product state are separated by an energy-barrier. Considering the time scale for such a reaction to take place, two times have to be considered.

One is the time required for the actual transformation like bond breaking or formation, or rearrangements. These occur at a time scale of picoseconds. It is independent of the temperature.

The second time is the waiting time of the system in the educt state for over-



coming the barrier. The average thermal energy  $kT$  at room temperature is only approximately 2.5 kJ/mol. Thus, the energy for overcoming larger barriers has to be provided by thermal fluctuations. They lead to a concentration of thermal energy in the vibration mode corresponding to the transition, which may enable the system to overcome the barrier. Following transition state theory it is possible to estimate the approximate waiting time  $T_w$  from the Arrhenius-equation:

$$T_w = \frac{1}{\nu} e^{\frac{E^\ddagger}{kT}} \quad (5.61)$$

where  $\nu$  is a characteristic frequency of the transition,  $E^\ddagger$  is the energy of the barrier,  $k$  is Boltzmann's constant and  $T$  is the temperature.

Examples of waiting times at different barrier heights are given in table 5.1. It is a rule of thumb that enzymatic reactions have a maximum barrier of 10 kcal/mol (42 kJ/mol). Higher barriers in a calculation indicate that the proposed mechanism does not explain the reaction.

However, in case of nitrogenase the total barrier of the reaction must not be compared with the total turnover rate. The reaction consists of many small steps, each of them occurs within less than a second. This time scale is limited by the electron transport, i.e. the dissociation of the two proteins. This electron transport consumes energy in form of MgATP and thus helps to overcome the total barrier in many small steps.

$E_A$ (kJ/mol)	time $T_w$
10	2 ps
20	120 ps
40	450 ns
60	1.6 ms
80	6 s
100	6 h

**Table 5.1** Waiting times for overcoming a reaction barrier for a vibration frequency of  $\nu = 3 \cdot 10^{13} \text{ s}^{-1}$  ( $\approx 1000 \text{ cm}^{-1}$ ) at 293.15 K.

## 5.5 Spin

Up to now I have treated electrons as particles which can fully be described by the three spatial coordinates  $x, y$ , and  $z$ . However, there are experiments, like the fine-structure in atomic spectra, which cannot be explained by this assumption. To solve this problem, Uhlenbeck and Goudsmith [136] proposed that the electron carries an intrinsic angular momentum, the spin.

### 5.5.1 Quantum mechanics

While the magnetic moment  $\vec{M}$ , associated with an angular momentum  $\vec{L}$  usually is

$$\vec{M} = \frac{\mu_B}{\hbar} \vec{L} = -\frac{e}{2m_e} \vec{L}, \quad (5.62)$$

the magnetic moment of the spin of an electron is

$$\vec{M}_s = 2\frac{\mu_B}{\hbar} \vec{S} = -\frac{e}{m_e} \vec{S}. \quad (5.63)$$

The so-called gyromagnetic ratio of the spin is twice that of the orbital angular momentum.<sup>9</sup>

The spin obeys the rules of a quantum mechanical angular momentum. While the total spin operator  $\mathcal{S}^2$  commutes with one of its components, usually chosen to be  $\mathcal{S}_z$ , the three components do not commute with each other. Thus only the total spin and its projection onto one axis in spin-space<sup>10</sup> may be calculated. A single electron is always in an eigenstate of  $\mathcal{S}^2$  with the eigenvalue  $\frac{3}{4}\hbar^2$ , thus its quantum number  $s = \frac{1}{2}$  with  $\langle \mathcal{S}^2 \rangle = \hbar^2 s(s+1)$ . Possible eigenvalues of  $\mathcal{S}_z$  are  $\pm\hbar\frac{1}{2}$ .

### 5.5.2 Pauli equation

In a fully relativistic description of the electrons, the Dirac-equation, each electron is described by a four component spinor wave function. Two of the components, the large components, describe the electronic part, while the other two, the small components, describe the positronic part always associated with the electron. However, the small components can be down-folded to the large components. The non-relativistic limit of the Dirac-equation is the Pauli-equation in which each electron is described by a two-component spinor wave function:

$$\left( \frac{(\vec{P} - e\vec{A})^2}{2m_e} + \frac{e}{m_e} \vec{S} \vec{B} + V \right) |\Psi\rangle = |\Psi\rangle \varepsilon \quad (5.64)$$

Here  $\vec{B}$  is the magnetic field and  $\vec{A}$  with  $\vec{B} = \vec{\nabla} \times \vec{A}$  is its vector potential. Relativistic effects include a modification of the kinetic-energy term as well as a term coupling the orbital angular momentum and the spin. While this so-called spin-orbit coupling is neglected in my DFT calculations, I include scalar relativistic corrections entering the kinetic-energy term.

<sup>9</sup>The more precise  $g_e$ -factor,  $g_e = 2\mu_e/\mu_B$ , obtained from quantum-electrodynamics is 2.00232...

<sup>10</sup>When neglecting spin-orbit coupling, spin-space and real space are decoupled.

As the wave function consists of two components,

$$\Psi(\vec{r}) = \begin{pmatrix} \Psi_1(\vec{r}) \\ \Psi_2(\vec{r}) \end{pmatrix} \quad (5.65)$$

all the operators occurring in the Pauli-equation, now replacing Schrödinger's equation, are two-by-two matrices.  $\vec{P}$  as well as  $V$  are diagonal as they act on each spin-direction equally.  $\vec{S}$ , the spin operator, has the form

$$\vec{S} = \frac{\hbar}{2} \vec{\sigma} \quad (5.66)$$

with  $\vec{\sigma}$  being the Pauli matrices

$$\sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \quad \sigma_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \quad \sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \quad (5.67)$$

The total density of one electronic state is still given by the operator  $n(\vec{r}) = |\vec{r}\rangle\langle\vec{r}|$  and the spin density  $n_s$  in  $x, y$ , and  $z$  direction are given by:

$$n(\vec{r}) = \Psi_1^*(\vec{r})\Psi_1(\vec{r}) + \Psi_2^*(\vec{r})\Psi_2(\vec{r}) \quad (5.68)$$

$$n_{sx}(\vec{r}) = \Psi_1^*(\vec{r})\Psi_2(\vec{r}) + \Psi_2^*(\vec{r})\Psi_1(\vec{r}) = 2\Re(\Psi_1^*(\vec{r})\Psi_2(\vec{r})) \quad (5.69)$$

$$n_{sy}(\vec{r}) = -i(\Psi_1^*(\vec{r})\Psi_2(\vec{r}) - \Psi_2^*(\vec{r})\Psi_1(\vec{r})) = 2\Im(\Psi_1^*(\vec{r})\Psi_2(\vec{r})) \quad (5.70)$$

$$n_{sz}(\vec{r}) = \Psi_1^*(\vec{r})\Psi_1(\vec{r}) - \Psi_2^*(\vec{r})\Psi_2(\vec{r}) \quad (5.71)$$

The magnetization density can be obtained from  $\vec{m} = -\frac{ge\hbar}{4m_e}\vec{n}_s$ . In case of a many-particle system, the total density and the spin densities are the sums of the corresponding one-particle densities of the occupied states.

In the special case of each electron of a given system being in an eigenstate of  $\mathcal{S}_z$  the system shows a collinear spin density. The eigenstates of  $\mathcal{S}_z$  are:

$$\Psi_{\uparrow}(\vec{r}) = \begin{pmatrix} \Psi_1(\vec{r}) \\ 0 \end{pmatrix} \quad \text{and} \quad \Psi_{\downarrow}(\vec{r}) = \begin{pmatrix} 0 \\ \Psi_2(\vec{r}) \end{pmatrix}. \quad (5.72)$$

Thus in this case, one of the spinor components of each electron vanishes, each electron may be described by a single spatial function. Correspondingly,  $n_{sx}$  and  $n_{sy}$  vanish everywhere in space. The Pauli-equation factorizes in two separate equations, coupled only via the external magnetic field.

### 5.5.3 Spin density functional theory

In usual density functional theory, see section 5.1, the ground state of a system can be characterized completely by the electron density  $n(\vec{r})$ . This theory is able to describe interacting electrons in a scalar (electrostatic) external potential. It

is sufficient for systems with equal numbers of spin-up and spin-down electrons in the absence of an external magnetic field. Apart from the density, the system can also be characterized completely by the effective Kohn-Sham potential  $v_{eff}$ .

For magnetic systems, the density functionals turned out to be of much simpler form if not only  $n(\vec{r})$  but also the magnetization density  $\vec{m}(\vec{r})$  are considered in the functional.

First, spin density functional theory has been applied to systems with uniaxial spin densities, thus with vanishing  $n_{sx}$  and  $n_{sy}$ . In these cases, each electron is either spin-up or spin-down. The Pauli-equation factorizes and the two spin-directions can be solved separately.

Later, also non-collinear systems, requiring two-component spinor wave functions, have been investigated [137–139].

Hobbes et al. [140] adapted the non-collinear description to the PAW method independently of our group. Our implementation differs from theirs in that we do allow for full non-collinearity also for the augmentation spin density.

#### 5.5.4 Spin contamination

$S_z$  is a one-particle operator, its expectation value can be calculated from the DFT wavefunction. In case of a collinear calculation, each electron is in an eigenstate of  $s_{i,z}$  and the total  $\langle S_z \rangle = \hbar/2(N_\uparrow - N_\downarrow)$  with  $N_\uparrow$  and  $N_\downarrow$  being the number of electrons with spin up and spin down, respectively.

The total spin operator,  $S^2$ , however, is a two-particle operator. *A priori* it is thus not well-defined how to obtain its expectation value from the non-interacting DFT reference wave function. This reference wave function has the same density and spin-density as the physical system but there is in principle no information on two-particle properties such as the total spin.

However, it is possible to apply the total spin operator  $S^2$  on an orthonormalized Slater-determinant like the Kohn-Sham wave function. With  $|i\rangle$  and  $|j\rangle$  being one-particle states, the expression is, as derived in appendix A.1:

$$\langle S^2 \rangle = \underbrace{\left( \frac{\hbar}{2} \int \vec{n}_s(\vec{r}) d^3r \right)^2}_{\text{classical}} + \underbrace{\frac{\hbar^2}{4} \sum_{ij} \left( 3\delta_{ij} - |\langle i_\uparrow | j_\uparrow \rangle - \langle i_\downarrow | j_\downarrow \rangle|^2 - 4|\langle i_\uparrow | j_\downarrow \rangle|^2 \right)}_{\text{quantum-mechanical}} \quad (5.73)$$

In collinear as well as in non-collinear calculations, both terms of (5.73) may contribute. In DFT the spins are treated classically in the sense that the total spin is not constrained to half-integer values. In case of non-collinear magnetism, not even  $S_z$  is constrained to half-integer values.

The first term in equation (5.73) is the total spin as it would be obtained from a classical spin- (or magnetization-) density. Its value is  $\hbar^2 \ell^2$  and its direction is

arbitrary as spin-space and real space are decoupled. This classical part is derived from the spin density, which is well-defined within spin density functional theory. The quantum mechanical part uses the one-particle states of the reference system which cannot be expected to lead to the same total spin as the real interacting system – not even with a perfect functional.

The DFT wave function is no spin eigenstate. This leads to a deviation of  $\langle S^2 \rangle$  from equation (5.73) from  $\hbar^2 \ell(\ell + 1)$  with half-integer  $\ell$ . This deviation is called spin contamination.

In the following, I describe occurrence of spin contamination in some examples.

For systems with singlet ground states and zero spin density (closed-shell systems), both parts of equation (5.73) vanish – the DFT wave function is a real singlet wave function.

Using the PBE functional described in section 5.1.4, spin contamination is small, whenever the spin density is positive everywhere in space.<sup>11</sup> This is the case for the nitrogen atom in its  $S = 3/2$  ground state, see table 5.2.

system	$\ell = \frac{1}{2}  \int \vec{n}_s $	$\ell(\ell + 1)$	$\langle S^2 \rangle / \hbar^2$
N	1.50000	3.75000	3.75252
$[\text{Fe}_2\text{S}_2]^{2+}$	0.00000	0.00000	4.01049

**Table 5.2** Spin contamination in various systems.

This situation changes, when one turns to systems that exhibit local spins but are experimentally found to be in a singlet state. The local spins can for example be caused by partially filled d-shells located at individual atoms. antiparallel alignment of these local spins can result in a singlet state.

Examples are the hydrogen molecule in its dissociation limit, antiferromagnetic solids, or the  $[\text{Fe}_2\text{S}_2]^{2+}$  cluster which exhibits two antiparallel  $S = 5/2$  iron sites. In these cases, two equal spins interact. For simplicity, I discuss a system of two interacting spins  $s_i = 1/2$  (like the dissociated  $\text{H}_2$  molecule). Larger spins show the same effects. The eigenstates of  $s_{z,i}$  are  $|\uparrow\uparrow\rangle$ ,  $|\uparrow\downarrow\rangle$ ,  $|\downarrow\uparrow\rangle$ , and  $|\downarrow\downarrow\rangle$ . The total spin operator  $S^2$ , however, has different eigenstates  $|\ell, m\rangle$ :

$$\begin{aligned}
 |0, 0\rangle &= 1/\sqrt{2}(|\uparrow\downarrow\rangle - |\downarrow\uparrow\rangle) \\
 |1, 0\rangle &= 1/\sqrt{2}(|\uparrow\downarrow\rangle + |\downarrow\uparrow\rangle) \\
 |1, 1\rangle &= |\uparrow\uparrow\rangle \\
 |1, -1\rangle &= |\downarrow\downarrow\rangle
 \end{aligned}$$

<sup>11</sup>In case of non-collinear calculations this means that  $n_s$  projected onto any direction has the same sign everywhere.

Thus, there is only one singlet state ( $|0,0\rangle$ ), but it cannot be described by a single determinant. Therefore the corresponding Kohn-Sham Slater determinant, which is a single determinant by definition, cannot be a singlet wave function. The integral over the spin density, and thus the classical  $\ell$

$$\ell = \frac{1}{2} \left| \int \vec{n}_s(\vec{r}) d^3r \right| \quad (5.74)$$

vanish, but  $\langle S^2 \rangle$  differs from zero. This can be seen in table 5.2 for the  $[\text{Fe}_2\text{S}_2]^{2+}$  system. Other antiferromagnetic systems behave similarly. The classical  $\ell$ , obtained from integrating the spin density is in agreement with the experimental total spin, while  $\langle S^2 \rangle$ , evaluated via equation (5.73) from the one-particle states, differs from it. Thus, I always calculate the total spin as

$$S = \ell = \frac{1}{2} \left| \int \vec{n}_s(\vec{r}) d^3r \right|. \quad (5.75)$$

A different problem arises, when truly non-collinear states are described. An example is the  $[\text{Fe}_3\text{S}_4]^+$  cluster, which will be elaborately discussed in section 6.4 on page 73. In this system, three spins  $S = 5/2$  interact. The experimental total spin is  $S = 1/2$ , which is also obtained by the simple but fully quantum-mechanical Heisenberg model described in the next chapter. In this model the spins align with angles of around  $120^\circ$  between them. In DFT the spins are treated classically: they are arranged with angles of nearly exactly  $120^\circ$  between them resulting in a total spin of  $S = 0$  which is not possible within quantum theory. In this case the state can be interpreted as a combination of two states with  $S = 1/2$  and  $m = \pm 1/2$ . However, in more complicated cases, such as the FeMoco of nitrogenase, such interpretations are not possible any more and the total spin of a real non-collinear system cannot be obtained from DFT calculations.

### 5.5.5 Collinear vs. non-collinear calculations

Here I discuss the advantages and disadvantages of collinear and non-collinear descriptions of the spin density in DFT calculations.

#### Non-collinear states

In quantum mechanics, any alignment of a certain spin with respect to some axis, lets say the  $z$ -axis, can be expressed by linear combinations of eigenstates of  $\mathcal{S}_z$ . Thus there is no need for non-collinear calculations, as eigenstates of  $\mathcal{S}_z$  correspond to collinear calculations.

DFT, however, is not able to correctly describe spin states if they are not in their highest or lowest possible  $m$ -value. This can easily be tested by calculating

an  $S = 2$  system like  $\text{Fe}^{2+}$  with  $S_z = 1$  or  $0$ . The energies are different,  $S_z = 2$  has the lowest energy. Of course,  $S_z = -2$  has the same energy as  $S_z = 2$ . Thus the local spin of any atom has to be aligned with the global  $z$ -axis to be correctly described in a collinear DFT calculation.

In systems with frustrated spins, systems with non-collinear spins, there are atoms with their local spin not aligned with the global  $z$ -axis. These systems need a two-component non-collinear description for being able to exhibit a local spin in a direction different from the global  $z$ -axis.

### Collinear states

Calculations which result in a collinear spin density also benefit from non-collinear formulations. The physically meaningful results, however, are the same as in collinear calculations. The main difference is that the results depend on the starting conditions in collinear calculations, while they are independent of the starting conditions in a non-collinear formulation.

The spin ordering<sup>12</sup> of a collinear calculation depends on the starting conditions of the calculation. If the spin of one atom starts in the up direction, a spin flip would have to proceed via a non-polarized transition state, which is energetically unfavorable. Thus spin flips usually do not take place and the spin ordering of the starting density is preserved. This may easily lead to metastable states, as the spin ordering is not known for complicated systems like the FeMo-cofactor. Moreover, the spin ordering frequently changes during chemical reactions. The barrier for a spin-flip of an  $\text{Fe}^{2+}$  site ( $S = 2$ ) can be approximated by calculating the same system (a one-center iron-sulfur cluster) with  $S_z = 0$ . This state is higher in energy by 108 kJ/mol than the system with  $S_z = 2$ . Such energy barriers are usually not overcome during an electronic structure relaxation. In non-collinear calculations, the spins may continuously rotate without changing their magnitude, which enables spin-flips.

Moreover, in collinear calculations with fixed occupations, the total  $S_z$  is fixed. This  $S_z$ , however, is not known for all states that have to be calculated. In a non-collinear calculation, the total  $S_z$  is not fixed but optimized during electronic and structural relaxation.

Of course, non-collinear calculations do not only have advantages over collinear ones: the computational effort is largely increased. In a collinear calculation, each electron is described by a single spatial function, and moreover, this one particle wavefunction may be chosen to be real. In a non-collinear calculation, each electron is described by two complex spatial functions. This increases the required memory and computation time. Table 5.3 summarizes the advantages and disadvantages of collinear vs. non-collinear calculations.

<sup>12</sup>the question which atoms exhibit spin-up and which exhibit spin-down

Collinear	Non-collinear
Frustrated states give incorrect results	Frustrated states described better
Metastable states	Metastable states avoided
10 calculations at known total spin	1 calculation
Eigenfunction of $\mathcal{S}_z$ , no eigenfunction of $\mathcal{S}^2$	No eigenfunction of $\mathcal{S}_z$ , no eigenfunction of $\mathcal{S}^2$
Per case:	
CPU-time: 1 Memory: 1	CPU-time: $\times 2.6$ Memory: $\times 3.9$

**Table 5.3** Comparison of collinear and non-collinear DFT calculations. The number of required cases to calculate is given for the resting state of FeMoco with the assumption of three-fold rotational symmetry.

Thus, a non-collinear calculation with collinear result can easily be verified by collinear calculations. On the other hand, it is not possible to obtain the global spin ordering and total spin from collinear calculations only. In principle all possible total spins, and for each of them all possible spin orderings, may be calculated collinear. This will provide the same ground state as obtained from one single non-collinear calculation. However, it is not known from collinear calculations that the ground state is indeed collinear. A non-collinear ground state will always be overseen by collinear calculations.



## 6 Discrete spins – the Heisenberg model

*Mathematics is a game played according to certain simple rules  
with meaningless marks on paper.*

David Hilbert

While continuous spin densities have been described in section 5.5.3 on page 57 the focus of this chapter will be discrete spins and their interactions. A high-spin ferric iron site has five electrons with parallel spins. Thus its net spin is  $S = 5/2$ . The spin of a ferrous iron site ( $\text{Fe}^{2+}$ ) is  $S = 2$ . While DFT treats the spins classically, a simple quantum mechanical description of the interaction of spins is provided by the Heisenberg spin Hamiltonian.

Operators are expressed in calligraphic letters ( $\mathcal{H}, \mathcal{S}$ ), quantum numbers of single sites in small letters ( $s, m$ ) and quantum numbers of the whole considered system in capital letters ( $S, M$ ).

### 6.1 Heisenberg spin Hamiltonian

Consider a system of fixed spins  $s_i$ . The general linear expression for the energy of interaction between the spins is

$$\mathcal{H} = \sum_{i>j} \vec{\mathcal{S}}_i \mathbf{A}_{ij} \vec{\mathcal{S}}_j. \quad (6.1)$$

with  $i$  and  $j$  labeling the atomic sites. The sum runs over all  $j$ , and for each  $j$  over all  $i > j$ .  $\mathbf{A}_{ij}$  is the coupling matrix.  $\mathcal{S}_i$  is the operator for the spin of site  $i$ , a three-dimensional vector.

For isotopic coupling the matrices  $\mathbf{A}_{ij}$  are reduced to equal diagonal elements resulting in the Heisenberg spin Hamiltonian

$$\mathcal{H} = \sum_{i>j} J_{ij} (\vec{\mathcal{S}}_i \vec{\mathcal{S}}_j) = \sum_{i>j} J_{ij} (\mathcal{S}_{x,i} \mathcal{S}_{x,j} + \mathcal{S}_{y,i} \mathcal{S}_{y,j} + \mathcal{S}_{z,i} \mathcal{S}_{z,j}). \quad (6.2)$$

$J$  is positive for antiferromagnetic coupling.

Using the definitions of the shift operators

$$\begin{aligned} \mathcal{S}_- &= \mathcal{S}_x - i\mathcal{S}_y \\ \mathcal{S}_+ &= \mathcal{S}_x + i\mathcal{S}_y = \mathcal{S}_-^\dagger, \end{aligned} \quad (6.3)$$

the scalar product in equation (6.2) can be rewritten to

$$\vec{S}_i \vec{S}_j = S_{z,i} S_{z,j} + \frac{1}{2} (S_{+,i} S_{-,j} + S_{-,i} S_{+,j}) \quad (6.4)$$

which can be used for calculations.

Equation (6.2) is sufficient for describing interactions between equal-oxidized transition metal atoms. If two atoms in different oxidation states interact, there could be resonance delocalization of (at least) one electron. In this way one  $\text{Fe}^{3+}$  and one  $\text{Fe}^{2+}$  give a pair of two  $\text{Fe}^{2.5+}$ . This interaction has first been described by Zener [141] and called *double exchange*.

Double exchange favors the spins of the interacting atoms to be parallel. Anderson and Hasegawa [142] expressed it as

$$\mathcal{H} = \sum_{i>j} J_{ij} (\vec{S}_i \vec{S}_j) + \sum_{i>j} -B_{ij} \left( \vec{S}_i \vec{S}_j + \frac{1}{2} \right). \quad (6.5)$$

In case of double exchange one electron partially resides on one site, partially on the other. This causes a weak chemical bond between these sites. The bonding character is only present in one spin direction as the other direction is filled. While the occurrence of such mixed-valence states plays an important role in the electronic structure of the ground state of nitrogenase, see section 9.2.1 on page 108 it will not be considered in more detail in this chapter.

Now the Hamiltonian is defined (6.2) a basis set is required to perform calculations on model systems.

### 6.1.1 Basis set

Each site  $i$  has the spin  $s_i$ , which is its first quantum number, describing the total spin, see equation (6.7). The projection of the spin onto the  $z$ -axis is given by the quantum number  $m$  (6.8). In the literature, this number is usually labeled  $m_s$  or  $s_z$ . I prefer  $m$  in order to avoid too many indices. The same algorithm is valid for angular momenta where these quantum numbers are usually labeled  $\ell$  and  $m$ . Thus the spin state of one site is referred by  $|s_i, m_i\rangle$ . The spin state of the whole cluster is the product state of its parts:

$$|s_1, m_1\rangle \otimes |s_2, m_2\rangle \otimes \dots = |s_1, m_1, s_2, m_2, \dots\rangle \quad (6.6)$$

$$\mathcal{S}_i^2 |\dots, s_i, m_i, \dots\rangle = |\dots, s_i, m_i, \dots\rangle \hbar^2 s_i(s_i + 1) \quad (6.7)$$

$$\mathcal{S}_{z,i} |\dots, s_i, m_i, \dots\rangle = |\dots, s_i, m_i, \dots\rangle \hbar m_i. \quad (6.8)$$

Since the  $s_i$  are not varied, the basis can be reduced to  $|m_1, m_2, m_3, \dots\rangle$ .  $N = (2s_1 + 1) \cdot (2s_2 + 1) \cdot (2s_3 + 1) \cdot \dots$  basis functions are required in order to completely describe the system.

## 6.2 The ground state of a spin-coupled system

Now that the Hamiltonian is set up using a proper basis set, the next task is to find its eigenstates. Diagonalization of the Hamilton-matrix, though rather involved, provides all eigenstates. It has been used to describe the three-site model in section 6.4 with 216 basis functions.

In contrast to that, the iterative method presented afterwards allows to treat much larger systems but only provides a few low-energy states. In larger systems like the FeMo cofactor, more than  $10^5$  basis functions may be required. Diagonalization of matrices of that size is computationally not feasible.

### 6.2.1 Diagonalization of the Hamilton-matrix

The Hamiltonian in the basis  $|m_1, m_2, m_3, \dots\rangle$  can be composed of the spin operators:

$$\langle m_1, m_2, \dots | \mathcal{S}_{z,i} | m'_1, m'_2, \dots \rangle = \hbar m'_i \delta(m_1, m'_1) \delta(m_2, m'_2) \dots \quad (6.9)$$

$$\begin{aligned} \langle m_1, m_2, \dots | \mathcal{S}_{+,1} | m'_1, m'_2, \dots \rangle &= \hbar \sqrt{s_1(s_1 + 1) - m'_1(m'_1 + 1)} \cdot \\ &\cdot \delta(m_1, m'_1 + 1) \delta(m_2, m'_2) \dots \end{aligned} \quad (6.10)$$

$$\begin{aligned} \langle m_1, m_2, \dots | \mathcal{S}_{-,1} | m'_1, m'_2, \dots \rangle &= \hbar \sqrt{s_1(s_1 + 1) - m'_1(m'_1 - 1)} \cdot \\ &\cdot \delta(m_1, m'_1 - 1) \delta(m_2, m'_2) \dots \end{aligned} \quad (6.11)$$

$$\mathcal{S}_{x,i} = \frac{1}{2}(\mathcal{S}_{+,i} + \mathcal{S}_{-,i}) \quad (6.12)$$

$$\mathcal{S}_{y,i} = \frac{1}{2i}(\mathcal{S}_{+,i} - \mathcal{S}_{-,i}) \quad (6.13)$$

with  $\delta(m, m')$  being the Kronecker symbol,  $\delta(m, m') = 1$  for  $m = m'$  and 0 otherwise. Knowing these matrices  $\mathcal{S}_{x,i}$ ,  $\mathcal{S}_{y,i}$ , and  $\mathcal{S}_{z,i}$  the total Hamiltonian can be set up as

$$\mathcal{H} = \sum_{i>j} J_{ij} (\mathcal{S}_{x,i} \mathcal{S}_{x,j} + \mathcal{S}_{y,i} \mathcal{S}_{y,j} + \mathcal{S}_{z,i} \mathcal{S}_{z,j}). \quad (6.14)$$

All  $\mathcal{S}$ -operators and the Hamiltonian  $\mathcal{H}$  are  $N \times N$  matrices. The indices labeling their entries will be  $n$  and  $k$ . They combine  $m_1, m_2, \dots$  into one number. Eigenstates of the Hamiltonian can be found by diagonalizing it, i. e. finding a unitary matrix  $\mathcal{U}$  transforming the Hamiltonian to a diagonal form:

$$\mathcal{U}_{in}^\dagger \mathcal{H}_{nk} \mathcal{U}_{kj} = \delta_{ij} \varepsilon_i \quad (6.15)$$

$$\mathcal{H} |n^H\rangle = |n^H\rangle \varepsilon_n \quad (6.16)$$

In this way it is possible to find the eigenstates  $|n^H\rangle$  of the Hamiltonian  $\mathcal{H}$ . The eigenstates are column-vectors of the unitary matrix  $\mathcal{U}$ . The eigenvalues reflect the

spin-coupling energy of the respective state. Most of the energy-eigenvalues are degenerate. Linear combinations of states within one multiplet are still eigenstates of the Hamiltonian.

The next task is to find the quantum numbers  $S$  and  $M$  of the eigenstates of the Hamiltonian. This is achieved by diagonalizing the matrix  $\langle n^H | \mathcal{S}^2 | k^H \rangle$  containing only states  $|n^H\rangle$  and  $|k^H\rangle$  within one multiplet. It leads to eigenstates  $|n^{H,S^2}\rangle$  of the Hamiltonian and the total spin operator  $\mathcal{S}^2$ :

$$\mathcal{S}^2 = (\mathcal{S}_{x,1} + \mathcal{S}_{x,2} + \dots)^2 + (\mathcal{S}_{y,1} + \mathcal{S}_{y,2} + \dots)^2 + \mathcal{S}_z^2 \quad (6.17)$$

These states exhibit a well-defined total spin  $S$ ,  $\langle n^{H,S^2} | \mathcal{S}^2 | n^{H,S^2} \rangle = \hbar^2 S(S+1)$ .

As there may still be states degenerate in both their energy and their total spin, the  $|n^{H,S^2}\rangle$  may be further characterized according to the  $z$  component of their total spin. This leads to states  $|n^{H,S^2,S_z}\rangle$ . It will be shown in section 6.4 that even this set may contain degenerate states.

### 6.2.2 Iterative search for the ground state

While the Hamilton matrix is small for model systems with few interacting sites, a model for the cofactor of nitrogenase would contain about 135 000 basis functions. The resulting matrix can not be diagonalized straight-forwardly. I am only interested in the ground state, and the first few excited states. A diagonalization also calculates all excited states.

I used an alternative approach to solve the electronic structure problem which is similar to the approach of Car and Parrinello [130]: damped classical mechanics.

The iteration is started with a random trial wave function  $|\Psi\rangle$ . The Hamiltonian provides the total energy of the system. The derivative of Schrödinger's equation with respect to  $\langle\Psi|$  corresponds to the classical force term. In order to perform classical mechanics, a fictitious mass  $m_\Psi$  is assigned to the wave function. This leads to the fictitious Lagrangian

$$\mathcal{L} = \langle\dot{\Psi}| m_\Psi |\dot{\Psi}\rangle - \langle\Psi| \mathcal{H} |\Psi\rangle + \text{constraints} \quad (6.18)$$

Using the Euler Lagrange equation, the following equation of motion can be derived:

$$m_\Psi |\ddot{\Psi}\rangle = -\frac{\partial E}{\partial \langle\Psi|} + \text{constraints} = -\mathcal{H} |\Psi\rangle + \text{constraints}. \quad (6.19)$$

Including a friction, one obtains:

$$m_\Psi |\ddot{\Psi}\rangle = -m_\Psi \alpha |\dot{\Psi}\rangle - \mathcal{H} |\Psi\rangle + \text{constraints} \quad (6.20)$$

The constraints are required to assure the normalization and, in case of multiple states, the orthogonality. This differential equation is of second order. It is

integrated numerically by discretization using the Verlet algorithm [129].

$$\dot{\Psi}(t) \approx \frac{\Psi(t + \Delta t) - \Psi(t - \Delta t)}{2\Delta t} \quad (6.21)$$

$$\ddot{\Psi}(t) \approx \frac{\Psi(t + \Delta t) - 2\Psi(t) + \Psi(t - \Delta t)}{(\Delta t)^2} \quad (6.22)$$

$$(6.23)$$

Thus, knowing two successive wave functions it is possible to estimate the next one:

$$\Psi(t + \Delta t) = \frac{1}{1 + a} \left[ 2\Psi(t) - (1 - a)\Psi(t - \Delta t) - \frac{\mathcal{H}\Psi(t)}{m_\Psi}(\Delta t)^2 \right] \quad (6.24)$$

The new friction-variable  $a$  is defined as  $a = \frac{\alpha\Delta t}{2}$ . In this formulation  $a$  can be varied from 0 (free motion) to 1 (steepest descent with a factor varying with  $m_\Psi$ ). In this way the groundstate can be calculated from random starting wave function coefficients.  $\Delta t$  and  $m_\Psi$  have to be chosen properly to lead to fast convergence. However, the converged result does not depend on these choices.

Up to now, I have omitted the constraints appearing in equation (6.19). In case of a single wavefunction, normalization can be achieved by simply normalizing the resulting function  $\Psi(t + \Delta t)$ . However, this violates energy conservation in undamped dynamics. Energy conservation is useful to test the correct implementation of the algorithm in the program. In the simulations, however, it is not necessary to use a normalization scheme that conserves the energy, because a friction is applied to the system, which does not conserve the energy either.

A different normalization scheme is to use the constraint  $\langle \Psi(t + \Delta t) | \Psi(t + \Delta t) \rangle = 1$  and the method of Lagrange parameters.

$$\Psi(t + \Delta t) = \underbrace{\frac{1}{1 + a} \left[ 2\Psi(t) - (1 - a)\Psi(t - \Delta t) - \frac{\mathcal{H}\Psi(t)}{m_\Psi}(\Delta t)^2 \right]}_{\Psi'} + \underbrace{\frac{2\Psi(t)(\Delta t)^2}{m_\Psi(1 + a)}}_{\Phi} \lambda \quad (6.25)$$

$$\Psi(t + \Delta t) = \Psi' + \Phi \lambda \quad (6.26)$$

The Lagrange parameter  $\lambda$  is obtained via

$$\langle \Psi(t + \Delta t) | \Psi(t + \Delta t) \rangle = 1 \quad (6.27)$$

$$\langle \Psi' | \Psi' \rangle + 2\langle \Psi' | \Phi \rangle \lambda + \langle \Phi | \Phi \rangle \lambda^2 = 1 \quad (6.28)$$

$$\lambda = -\frac{\langle \Psi' | \Phi \rangle}{\langle \Phi | \Phi \rangle} \pm \sqrt{\left( \frac{\langle \Psi' | \Phi \rangle}{\langle \Phi | \Phi \rangle} \right)^2 - \frac{\langle \Psi' | \Psi' \rangle - 1}{\langle \Phi | \Phi \rangle}} \quad (6.29)$$

This scheme is energy conserving and converges faster than the above mentioned one with direct normalization. However, the estimation of the Lagrange parameter requires additional computational effort.

As the ground state is in some cases degenerate, it is of interest to calculate a few additional states. These could also be constrained to be orthogonal to each other by Lagrange parameters, which allows to calculate more than one state simultaneously. However, it is not known beforehand how many states are degenerate and thus I decided to calculate them one by one. After convergence of the ground state, the second state  $|\Psi'_2\rangle$  is again chosen randomly. It is restricted to be orthogonal to the first state by transforming it via

$$|\Psi_2\rangle = |\Psi'_2\rangle - \langle\Psi_1|\Psi'_2\rangle |\Psi_1\rangle \quad (6.30)$$

after each time step. Next,  $|\Psi_2\rangle$  has to be normalized. In this way a series of states is calculated and orthogonalized to all previous states until the energy of one state lies significantly above that of the first state. Now the set of ground states is known and can be decomposed into eigenstates of  $\mathcal{S}^2$  and  $\mathcal{S}_z$ .

### Example

The iron-sulfur cluster  $[\text{Fe}_2\text{S}_2]^{2+}$  can be modeled as two interacting spins with  $s_1 = s_2 = 5/2$ . As the coupling is experimentally known to be antiferromagnetic the only coupling constant  $J_{12}$  is chosen positive. The actual value of this constant is unimportant at this stage of analysis because the interest is focused on the spin configuration of the ground state and not on excitation energies.

Knowing the fact that the coupling is antiferromagnetic one would suggest  $|m_1, m_2\rangle = |\frac{5}{2}, -\frac{5}{2}\rangle$  as ground state, meaning one atom completely spin-up and the other one spin-down. However, the result of the Heisenberg model is

$$|0^{H, S^2, S_z}\rangle = \frac{1}{\sqrt{6}} \left( -|-\frac{5}{2}, \frac{5}{2}\rangle + |-\frac{3}{2}, \frac{3}{2}\rangle - |-\frac{1}{2}, \frac{1}{2}\rangle + |\frac{1}{2}, -\frac{1}{2}\rangle - |\frac{3}{2}, -\frac{3}{2}\rangle + |\frac{5}{2}, -\frac{5}{2}\rangle \right)$$

This is a non-degenerate singlet wave function ( $S = 0$ ,  $S_z = 0$ ), while  $|\frac{5}{2}, -\frac{5}{2}\rangle$  would not be singlet.

The highest states exhibit  $S = 5$  as expected. In contrast to the singlet ground state, however, some of them are eigenstates of  $\mathcal{S}_{z,i}$  like  $|\frac{5}{2}, \frac{5}{2}\rangle$ .

### 6.2.3 Analysis

The resulting wave functions are eigenfunctions of  $\mathcal{S}^2$  and  $\mathcal{S}_z$  but not of  $\mathcal{S}_{z,i}$ , the  $z$  component of the spin on a single site. The expectation value of  $\mathcal{S}_{z,i}$ , required

for the spin projection coefficients which will be used in section 6.4 for calculating magnetic hyperfine parameters, can be calculated as

$$K_i = \frac{\langle \mathcal{S}_{z,i} \rangle}{S_z}. \quad (6.31)$$

The expectation values  $\mathcal{S}_{x,i}$  and  $\mathcal{S}_{y,i}$  vanish due to the uncertainty principle. The question if a spin ordering is collinear can be addressed using a different quantity, the angle between two spins defined via

$$\cos \alpha = \frac{\langle \mathcal{S}_1 \mathcal{S}_2 \rangle}{|\mathcal{S}_1| |\mathcal{S}_2|} \quad (6.32)$$

with  $|\mathcal{S}_1| = \sqrt{\langle \mathcal{S}_1^2 \rangle}$ . The length of the spin vectors are constant as the resulting wave functions still are eigenfunctions of  $\mathcal{S}_i^2$ .

### Example

In the example mentioned above with two  $s_1 = s_2 = 5/2$  the ground state is an eigenstate of  $\mathcal{S}_1 \mathcal{S}_2$ . The eigenvalue of the angle between the spins is  $180^\circ$ , thus the spins are really antiparallel. Note that, although the angle is well-defined,  $\mathcal{S}_{z,i}$  vanishes for each site.

It is interesting to note that the expectation value of the angle between the two spins in the high-spin state  $|\frac{5}{2}, \frac{5}{2}\rangle$  does not vanish. It is  $44.4^\circ$ .

In case of two interacting spins with  $s_1 = s_2 = 1/2$  there are four resulting states. In the singlet state, the wave function is an eigenfunction of the angle between the spins. The eigenvalue is  $180^\circ$ . In contrast to that, the wave functions of the three triplet states are no eigenfunctions of the angle between the spins. The expectation values are  $70.5^\circ$  ( $\cos \alpha = \frac{1}{3}$ ).

## 6.3 Classical spins – the classical Heisenberg model

In non-collinear DFT calculations, the spins are treated classically. Thus the wave function neither an eigenstate of  $\mathcal{S}_z$  nor of  $\mathcal{S}^2$ . The classical Heisenberg model may lead to a quantum-mechanically interpretation of the results from DFT. Spin vectors obtained from DFT are used to compute the corresponding set of spin coupling constants  $J_{ij}$ . These parameterize a quantum mechanical Heisenberg model, which allows physical interpretation.

In order to calculate  $J_{ij}$  from a given set of  $\vec{S}_i$ , it is necessary to calculate  $\vec{S}_i$  from a given set of  $J_{ij}$ .

The classical Heisenberg model uses spins that behave like classical angular momentum vectors. In contrast to the quantum mechanical model, all three components of each spin can be calculated simultaneously. The Hamiltonian is:

$$\mathcal{H} = \sum_{ij} J_{ij} \vec{S}_i \cdot \vec{S}_j \quad (6.33)$$

The sum over all  $i$  and  $j$  can be used easier than the one over  $i \neq j$  and gives the same result concerning the spin directions.

### Calculating $\mathbf{S}_i$ from given $\mathbf{J}_{ij}$

The fictious Lagrangian of the classical spin-system is obtained by introducing a fictious mass ( $m$ ) for spin-rotations

$$\mathcal{L} = \frac{1}{2} \sum_i m \dot{\vec{S}}_i^2 - \sum_{ij} J_{ij} \vec{S}_i \cdot \vec{S}_j + \sum_i \lambda_i (\vec{S}_i^2 - 1) \quad (6.34)$$

The lengths of the spin-vectors have to be preserved which is assured by the Lagrange parameters  $\lambda_i$ . Here I use unity as length of all spin vectors. They and the corresponding  $J$ -values may be scaled after the calculation. The derivations with respect to  $\vec{S}_i$  and  $\lambda_i$  yield

$$m \ddot{\vec{S}}_i = -2 \sum_j J_{ij} \vec{S}_j + 2 \vec{S}_i \lambda_i \quad (6.35)$$

$$\vec{S}_i^2 = 1 \quad (6.36)$$

leading to the discretized equations of motions with a friction term:

$$\begin{aligned} \frac{m}{(\Delta t)^2} (\vec{S}_i(t + \Delta t) - 2\vec{S}_i(t) + \vec{S}_i(t - \Delta t)) = \\ -2 \sum_j J_{ij} \vec{S}_j(t) + 2 \vec{S}_i(t) \lambda_i - m \alpha \frac{\vec{S}_i(t + \Delta t) - \vec{S}_i(t - \Delta t)}{2\Delta t}. \end{aligned} \quad (6.37)$$

Using  $a = \frac{\alpha \Delta t}{2}$  leads to the same expression for the next time step as in equation (6.26) with according Lagrange parameters.

These calculations can be performed very fast and the dynamic converges quite well. In this way the spin pattern can be calculated from a set of coupling parameters  $J_{ij}$ .



### Calculating $J_{ij}$ from given $S_i$

The estimation of spin coupling parameters  $J_{ij}$  from given spins  $\vec{S}_i$  can be used to physically interpret the spin coupling scheme obtained from DFT calculations.

While previous studies used a set of collinear calculations with different spin orderings to estimate  $J_{ij}$  [143, 144], I only use one non-collinear calculation.

I used an iterative method for obtaining a set of  $J_{ij}$  which have the given  $\vec{S}_i$  as ground state. Starting with properly chosen  $J_{ij}$  it is possible to calculate initial spin vectors  $S_i^0$ . In general, they differ from the given set  $\vec{S}_i$ .

In the following, I will change the notation in order to get more readable equations. Now the index  $i$  of  $S_i$  runs over  $3 \times N$  spatial coordinates. Moreover, the matrix  $J_{ij}$  is converted into a vector  $J_\alpha$ . As  $J_{ij}$  is symmetric with vanishing diagonal elements, the matrix as well as the vector have  $N(N-1)/2$  independent components.

Starting from  $S_i^0$  and using an approach similar to perturbation theory, the given final spins  $S_i^{(f)}$  are linearly approximated:

$$S_i^0 + \sum_{\alpha} \Delta J_{\alpha} \frac{\partial S_i}{\partial J_{\alpha}} = S_i^{(f)} \quad (6.38)$$

$\Delta J_{\alpha}$  has to be calculated. This is achieved by minimizing the difference between the left-hand side of equation (6.38) and the spins  $S_i^{(f)}$  obtained from DFT:

$$F[\Delta J_{\alpha}] = \sum_i \left[ S_i^0 + \sum_{\alpha} \Delta J_{\alpha} \frac{\partial S_i}{\partial J_{\alpha}} - S_i^{(f)} \right]^2 \rightarrow \min \quad (6.39)$$

$$\begin{aligned} F[\Delta J_{\alpha}] &= \underbrace{\sum_i (S_i^0 - S_i^{(f)})^2}_C + 2 \sum_{\alpha} \Delta J_{\alpha} \underbrace{\sum_i \frac{\partial S_i^0}{\partial J_{\alpha}} (S_i^0 - S_i^{(f)})}_{B_{\alpha}} \\ &+ \sum_{\alpha, \alpha'} \Delta J_{\alpha'} \underbrace{\sum_i \frac{\partial S_i^0}{\partial J_{\alpha}} \frac{\partial S_i^0}{\partial J_{\alpha'}}}_{A_{\alpha, \alpha'}} \Delta J_{\alpha} \end{aligned} \quad (6.40)$$

$$F[\Delta J_{\alpha}] = C + 2 \sum_{\alpha} \Delta J_{\alpha} B_{\alpha} + \sum_{\alpha, \alpha'} \Delta J_{\alpha'} A_{\alpha, \alpha'} \Delta J_{\alpha} \rightarrow \min \quad (6.41)$$

$$\frac{\partial F}{\partial \Delta J_{\alpha}} = 2B_{\alpha} + 2 \sum_{\alpha'} \Delta J_{\alpha'} A_{\alpha, \alpha'} = 0 \quad (6.42)$$

or, as matrix equation:

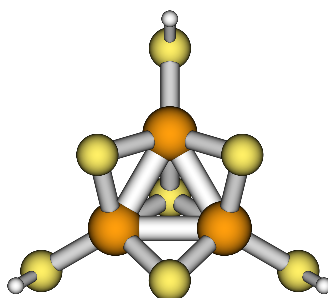
$$B + A \Delta J = 0 \quad \Delta J = -A^{-1} B \quad (6.43)$$

Changing the  $J_\alpha$  ( $= J_{ij}$ ) along the path provided by  $\Delta J$  iteratively leads to a set of  $J_{ij}$  which gives the same spins as the DFT calculation.

The derivative  $\frac{\partial S_i^0}{\partial J_\alpha}$  is calculated numerically as differential quotient. For that purpose, a unique spin-pattern out of a given set of  $J_{ij}$  it is necessary. This is achieved by freezing the three rotational degrees of freedom.

The algorithm only converges for reasonable starting conditions. It is advisable to choose large starting values for  $J_{ij}$  coupling spins with a large angle between them and small values for spins is larger distance or with small angle. Moreover, the converged results sometimes depend on the starting conditions. Thus there is in general no unique set of  $J_{ij}$  that can be obtained from given  $\vec{S}_i$ . In general, however, different sets of  $J_{ij}$  obtained from the same spin coupling pattern lead to the same ground state in the quantum mechanical Heisenberg model. All further analysis is based on this quantum mechanical ground state.

## 6.4 Application: three interacting spins in $[\text{Fe}_3\text{S}_4]^+$



**Figure 6.1** Structure of the  $[\text{Fe}_3\text{S}_4]^+$  complex. In this direction the approximate three-fold rotational symmetry is clearly visible. The cluster is connected via three sulfur bridges to cysteine residues of the protein which have been truncated here.

The iron sulfur cluster  $[\text{Fe}_3\text{S}_4]$  occurs in several ferredoxins, hydrogenases, and in aconitase. It can adopt at least two redox levels, namely  $[\text{Fe}_3\text{S}_4]^+$  and  $[\text{Fe}_3\text{S}_4]^0$  at biological conditions. Thus it acts as the functional unit in the redox enzymes. Here I will discuss the oxidized state  $[\text{Fe}_3\text{S}_4]^+$ , which exhibits three high-spin ferric iron sites.

The structure of the considered complex is shown in Fig. 6.1. It shows a cuboidal arrangement of iron and sulfur atoms with one iron atom removed. The remaining iron sites form a triangle with one sulfur atom below the center of the triangle and three sulfur atoms above the edges. The iron sites are connected to the protein via cysteine sulfur bridges. Thus all iron atoms are in an approximate tetrahedral environment while the whole structure shows three-fold rotational symmetry. The structure, EXAFS data [145], as well as measured Mössbauer isomer shifts and electric field gradients [146–149] indicate that the system is highly symmetric.

This symmetry seems to be broken in the spin structure. This can be seen from the hyperfine parameters listed in table 6.1 on the next page. They have been measured in  $[\text{Fe}_3\text{S}_4]^+$  clusters in various enzymes of different species. While the total spin of the cluster in all 7 of these species is  $S = 1/2$ , the individual site contributions differ between the species. They also differ for the three iron sites within each species.

In the following this discrepancy between the highly symmetric atomic structure and the virtually asymmetric spin structure will be explained by a non-collinear spin structure. The latter, even though being three-fold rotational symmetric, leads to highly asymmetric hyperfine parameters.

enzyme	$A_{iso}(\text{Fe}^{3+})$ (MHz)		
1 <i>Pyrococcus furiosus</i> ferredoxin [150]	−37.5	+26	−11
2 <i>Desulfovibrio gigas</i> hydrogenase [151] form 1	−44	+20	+3
3 <i>Desulfovibrio gigas</i> hydrogenase [151] form 2	−39	+23	
4 <i>Azotobacter vinelandii</i> ferredoxin I [152]	−41.9	+17.4	+2
5 <i>Azotobacter vinelandii</i> ferredoxin I [146]	−41	+18	±5
6 beef heart aconitase [153]	−30.6	+28.1	−10.5
7 <i>Chromatium vinosum</i> hydrogenase [154]	−44	+12	

**Table 6.1** Isotropic magnetic hyperfine parameters of  $[\text{Fe}_3\text{S}_4]^+$  complexes in various enzymes.

### 6.4.1 The Heisenberg model

The spin structure of the system can best be modeled with three discrete spins, each with  $s_i = 5/2$ . Kent et al. [146] pointed out that the total spin of  $S = 1/2$  in this system can only be achieved by two spin eigenstates. Linear combinations of these 2 states also result in  $S = 1/2$ . These two states result from the coupling of one spin with  $s_1 = 5/2$  with a combination of the other two spins to  $s_{23} = 2$  or  $s_{23} = 3$ . Using the nomenclature  $|s_{23}, s_1, S\rangle$  they are labeled  $|\Psi_1\rangle = |2, \frac{5}{2}, \frac{1}{2}\rangle$  and  $|\Psi_2\rangle = |3, \frac{5}{2}, \frac{1}{2}\rangle$ .

I investigated this system using the Heisenberg spin coupling scheme with three equal or similar  $J$  coupling constants following from the approximate symmetry of the system. Using equal coupling constants results in  $|\Psi_1\rangle$  and  $|\Psi_2\rangle$  as degenerate ground states of the system. Small differences in the  $J$  values remove the degeneracy. If  $J_{23} > J_{12} = J_{13}$  the pure state  $|\Psi_1\rangle$  is the ground state [146, 150], if  $J_{23} < J_{12} = J_{13}$  it is the pure state  $|\Psi_2\rangle$ . In all other cases the ground state is a mixture, a linear combination, of  $|\Psi_1\rangle$  and  $|\Psi_2\rangle$ .

While Kent et al. used the linear combination  $|\Psi\rangle = |\Psi_1\rangle\sqrt{1-\alpha^2} + |\Psi_2\rangle\alpha$  I feel that

$$|\Psi\rangle = |\Psi_1\rangle \cos \phi + |\Psi_2\rangle \sin \phi \quad (6.44)$$

is more convenient to use. The conversion is therefore  $\alpha = \sin \phi$ . Only states within  $0 \leq \phi \leq \frac{\pi}{6}$  are relevant because values outside that range would correspond to relabeling of the states. The state with  $\phi = \frac{\pi}{6}$  has  $s_{13} = 3$  and will be the ground state if  $J_{13} < J_{12} = J_{23}$ .

Diagonalization of the Heisenberg Hamiltonian provides the states  $|\Psi_1\rangle$  and  $|\Psi_2\rangle$  in the representation of  $|m_1, m_2, m_3\rangle$ <sup>1</sup>. They are required to calculate the hyperfine parameters of the spin-coupled system.

<sup>1</sup>Apart from solving the Heisenberg model, the states  $|\Psi_1\rangle$  and  $|\Psi_2\rangle$  can also be obtained using the Clebsch-Gordan coefficients

The intrinsic hyperfine parameter  $a_i$  of an atom depends on the spin density at its nuclear site. One of the contributions to  $a_i$  is the Fermi-contact term. In case of a single isolated site, the measured hyperfine parameter  $A_i$  is equal to  $a_i$ .

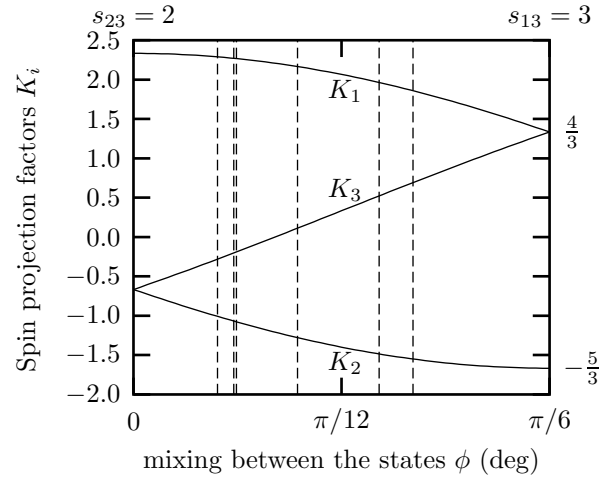
This situation changes in a system with several spin-coupled nuclei. The measured hyperfine parameter  $A_i$  of a site  $i$  in the spin coupled system depends on  $a_i$  and the projection of the local spin of site  $i$  on the total spin of the system [29, 155, 156]:

$$A_i = K_i a_i \quad \text{with} \quad K_i = \frac{\langle s_{z,i} \rangle}{\langle S_z \rangle}. \quad (6.45)$$

In an isolated site, the spin projection factor  $K_i$  becomes unity and  $A_i = a_i$ . In a coupled system, the measured hyperfine parameter decreases with increasing angle between the local spin of the site and the total spin of the system. In case of antiparallel alignment between the local and total spin, the sign of the measured hyperfine parameter is reversed compared to  $a_i$ .

In order to characterize a system in terms of the Heisenberg model and to compare the model with experimental results,  $a_i$  and  $\phi$ , the mixing parameter between the states  $|\Psi_1\rangle$  and  $|\Psi_2\rangle$  have to be known. They can be calculated for each cluster, as shown in the following.

The intrinsic isotropic hyperfine parameter  $a_i$  depends on the oxidation state and the chemical environment of the iron site but not on the spin coupling to the other sites. Thus it can be assumed that this  $a_i$  is the same for all three iron sites of one cluster.



**Figure 6.2** Spin projection factors in  $[\text{Fe}_3\text{S}_4]^+$ . The dashed lines represent the mixing of the states found for different enzymes.

Using  $\phi$  and  $|\Psi_1\rangle$  and  $|\Psi_2\rangle$  it is possible to calculate the spin projection co-

efficients  $K_i$  from equation (6.45). They are shown as solid lines in Fig. 6.2 on the preceding page. The differences between the three  $K_i$  for the same system explain the huge differences between the three measured hyperfine parameters of each  $[\text{Fe}_3\text{S}_4]^+$  cluster.

I used the experimental hyperfine parameters for the three iron sites of each of the seven systems, given in table 6.1 on page 74.  $\phi$  and  $a_i$  have been fitted to these three values. The obtained mixing parameters  $\phi$  for the seven systems are indicated in Fig. 6.2 on the preceding page as dashed lines. The results for  $\phi$  and  $a_i$  are given in table 6.2.

	$\sin^2 \phi$	$a_i$ (MHz)	$K_i$		
1	9.2 %	−18.6	1.963	−1.486	0.523
2	1.7 %	−17.9	2.267	−1.077	−0.190
3	4.2 %	−18.0	2.165	−1.277	0.112
4	1.6 %	−18.0	2.270	−1.068	−0.202
5	1.1 %	−17.8	2.289	−1.008	−0.281
6	11.9 %	−17.0	1.859	−1.550	0.691
7	0.0 %	−18.8	2.333	−0.667	−0.667

	theory			experiment		
	$A_{iso}$ (MHz)			$A_{iso}$ (MHz)		
1	−36.6	27.7	−9.7	−37.5	+26	−11
2	−40.6	19.3	3.4	−44	+20	+3
3	−39.0	23.0	−2.0	−39	+23	
4	−40.9	19.2	3.6	−41.9	+17.4	+2
5	−40.9	18.0	5.0	−41	+18	±5
6	−31.6	26.3	−11.7	−30.6	+28.1	−10.5
7	−43.8	12.5	12.5	−44	+12	

**Table 6.2** Results of the Heisenberg model for the systems 1–7. Top: the fitted parameters  $\phi$  and  $a_i$  and the spin projection coefficients obtained from  $\phi$ . Bottom: theoretical and experimental hyperfine parameters. The mixing parameter  $\phi$  is given in the form of the percentage of the  $|\Psi_2\rangle$  state in the ground state,  $\sin^2 \phi$ .

It can be seen that all the systems can be described well with this model. Considering the experimental uncertainty of about  $\pm 3$  MHz all experimental hyperfine parameters can be reproduced well.

For the systems 3 and 7, the hyperfine coupling of one iron site could not be measured by experiment. I used the remaining two values to estimate  $a_i$  and  $A_3$ . In case of system 3 this value is rather small (−2 MHz) and thus difficult to detect experimentally. In system 7,  $A_2$  and  $A_3$  are degenerate. This system is a special case, as the authors of the respective experimental study [154] point out that there

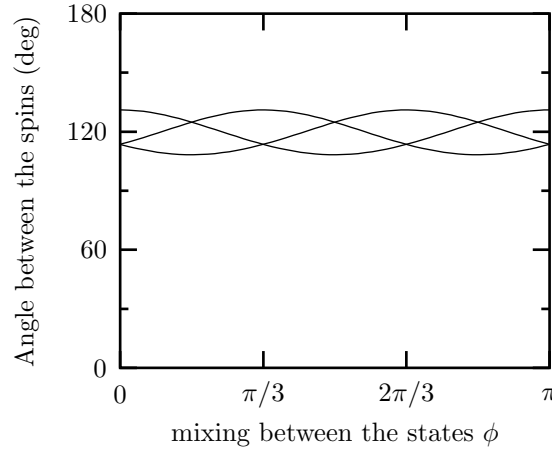
may be a fourth iron site interacting with the measured three.

It can be seen from table 6.2 that the mixing between the states  $|\Psi_1\rangle$  and  $|\Psi_2\rangle$  varies significantly between the considered systems. These variations are caused by slightly different coupling constants  $J$  in the systems. These coupling constants depend on structural deviations.

The intrinsic hyperfine parameter  $a_i$  is preserved quite well over the range of investigated systems. This shows the chemical similarity between the seven systems. It also justifies the assumption that  $a_i$  is the same for the three iron sites of one complex. The mean value of the intrinsic hyperfine parameter is  $a_i = 18.0$  MHz.

### 6.4.2 Non-collinear magnetism

Now that I have shown that the results of the Heisenberg model properly describe the spin coupling in  $[\text{Fe}_3\text{S}_4]^+$  it is legitimate to draw further conclusions from this model. While the hyperfine parameters have already been discussed in the literature cited above, it is also possible to extract information about the angles between the spins from the Heisenberg model.



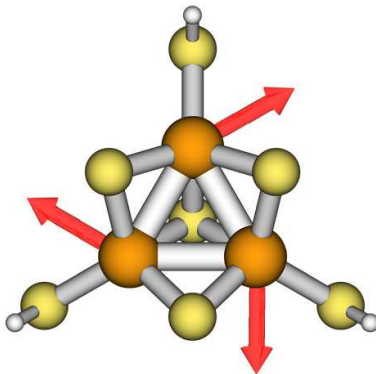
**Figure 6.3** Expectation values of the angles between the three spins of the  $[\text{Fe}_3\text{S}_4]^+$  cluster derived from the contributing states  $|\Psi_1\rangle$  and  $|\Psi_2\rangle$ .

Using the states  $|\Psi_1\rangle$  and  $|\Psi_2\rangle$  in the representation of the spin eigenstates of each site ( $|m_1, m_2, m_3\rangle$ ) it is straightforward to derive the expectation values of the angles between them according to equation (6.32). The result is depicted in Fig. 6.3. Depending on the mixture of the states, the angles vary around  $120^\circ$ <sup>2</sup>. In case of maxima or minima the respective angle is an eigenvalue while it is only

<sup>2</sup>between  $108.3^\circ$  and  $131.1^\circ$

an expectation value otherwise. The approximate threefold rotational symmetry of the atomic structure is preserved in the spin structure.

These angles of around  $120^\circ$  between the three spins are a clear indication for non-collinear spin ordering in that system.



**Figure 6.4** Illustration of the spins of  $[\text{Fe}_3\text{S}_4]^+$  as obtained from DFT calculations.

This finding has to be considered when describing the system with DFT. In non-collinear calculations the spins are treated classically, the wave function is neither an eigenstate of  $S_z$  nor of  $S^2$ . The antiferromagnetic coupling causes the spins to be in one plane with angles of  $120^\circ$  to each other. Thus the threefold symmetry is preserved and the results agree quite well with experiments. Also the intrinsic hyperfine parameters obtained from the calculations are the same for the three iron sites.

Collinear calculations restrict the system to eigenstates of  $S_z$ . There is no proper choice of  $S_z$  which takes care of the three-fold rotational symmetry of the system except  $S_z = 15/2$ , the high-spin state. This state does not represent the chemical behavior of the cluster, which shows antiferromagnetic coupling. If the experimental spin  $S = 1/2$  is chosen, one iron comes out to be spin-up, another one spin-down and the third iron shows a rather small spin. This deteriorates the three-fold symmetry as it influences density-related observables such as Mössbauer isomer shifts and electric field gradients.

Thus a model with discrete spins, the quantum mechanical Heisenberg model, shows that a treatment of the spin density allowing for non-collinearity is necessary for describing the continuous system with DFT.



# **Part III**

## **Application**



## 7 Considerations concerning the structure of nitrogenase

*Thus, the task is, not so much to see what no one has yet seen;  
but to think what nobody has yet thought, about that which everybody sees.*

Erwin Schrödinger

The cofactor of nitrogenase is embedded in the MoFe protein. It is essential to know the abilities of the protein to transport small particles to the cofactor in order to understand the reaction mechanism. The focus lies on protons, electrons,  $\text{N}_2$ , and  $\text{NH}_3$ .

Here, I will show how protons reach the cofactor and where they reach it. They move in special proton transport channels. Such channels can not be found directly for  $\text{N}_2$  or  $\text{NH}_3$  as both show only weak interaction with the protein. Electrons, another ingredient in the conversion cycle, are delivered by the Fe-protein and reach the FeMoco via the P-cluster as discussed in section 3.1.

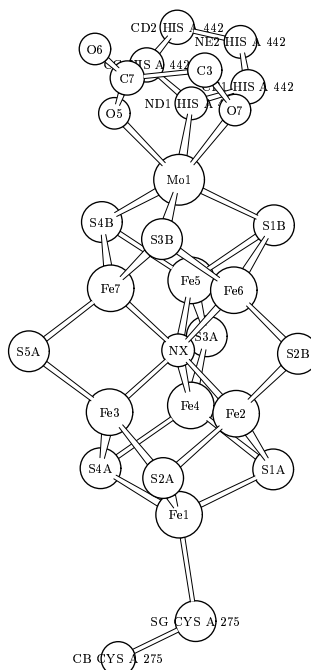
The investigations presented in this chapter are not based on DFT calculations but on an analysis of the crystal structure. I used the most accurate structure determination of the MoFe-protein published up to now, the 1M1N PDB data set, refined to 1.16 Å [19].

### 7.1 Proton channels

Here I will show that there is only one proton channel capable of transporting a series of protons from the surrounding solution to the cofactor. It starts at a cavity between the two halves of the MoFe-protein and proceeds via 14 steps to a water molecule near the iron site Fe7.

#### 7.1.1 The proton shuttle mechanism

Protons can be transported in aqueous or similar solutions using an effective mechanism of proton-hopping within hydrogen bonds followed by rotation of the proton acceptors. This is the reason for the high electric conductivity of acids or bases. The mechanism of hopping and rotation is illustrated in Fig. 7.2 and can be explained as follows:

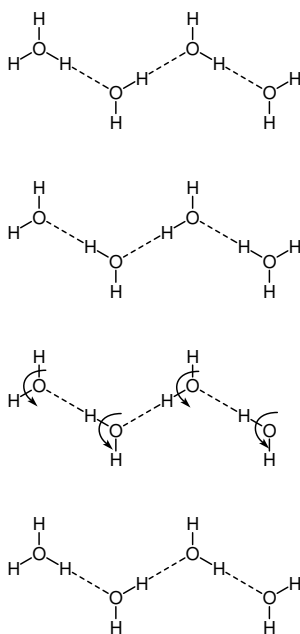


**Figure 7.1** Nitrogenase FeMoco with structure and nomenclature of the 1M1N [19] PDB data base.

1. In liquid water or any other protic system, the molecules are connected by hydrogen bonds. Now I consider a chain of water molecules which should transfer protons from one end to the other. Protons can change their partners within the hydrogen bond. This step transfers one proton from one end of the chain to the other without significant structural rearrangements of any atom.
2. A second proton transfer in the same direction requires prior rotation of the molecules. During this rotation, hydrogen bonds are broken and others are formed.
3. After a rotation of the molecules, the next proton is transferred.

Thus, in order to effectively transport protons, the transport units, acting simultaneously as acceptors and donors, have to be able to rotate at least around one axis. In a protein there are different types of transport units:

- Water molecules: while ordered water molecules are detected by X-ray diffraction, disordered molecules are invisible to diffraction methods. Disordered water molecules result in cavities within the experimental protein structure.



**Figure 7.2** Mechanism of proton transport by hopping and rotation. Protons are transported via the chain of water molecules from the left side to the right side.

- Hydrophilic groups of protein residues, i.e. the groups  $-\text{OH}$ ,  $-\text{NH}_2$ ,  $-\text{SH}$ . Amino acids offering these groups are listed in table 7.1.

This proton shuttle mechanism allows proteins to transport protons in channels filled with such transport units. Channels like that have already been found in the nitrogenase MoFe-protein [65, 68].

One proton transport channel has been described to lead, via water molecules, from the inner surface of the protein to the water pool around homocitrate. The inner surface is the region where the two halves of the protein are connected. In this region there are large pockets filled with water [65, 68].

Another proton transport channel [65, 68] connects the surface to the sulfur bridge S2B. The last transport unit of this channel is the imidazole ring of His $\alpha$ 159. This channel is only able to transport one proton to the cofactor, because a second proton would require the whole imidazole ring of histidine to rotate, which is unlikely to happen. The proton has to be transferred back to histidine. This channel will be important for explaining the catalytic cycle as described in section 10.1.2. However, as it is not able to transport more than one proton, it will not be considered here.

The new results of my work are (1) that I found three more water molecules which extend the firstly mentioned channel is extended from the water pool around the homocitrate to one bridging sulfur atom of the cofactor, and (2) a verification that this channel is the only one to the cofactor in the most accurate X-ray structure available. The previous studies [65, 68] have been based on the less accurate 2MIN X-ray structure. Moreover, only water molecules have been considered as proton transport units there.

### 7.1.2 Systematic search for proton channels – calculational details

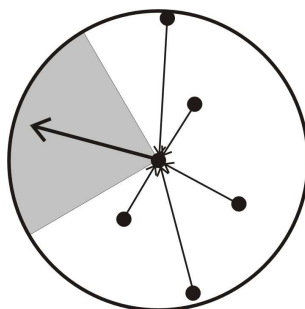
While previous studies [65, 68] considered the less accurate 2MIN PDB dataset and restricted their search for proton channels to water molecules detected by diffraction, my approach is somewhat more general: (1) I use water molecules and (2) the amino acid residues listed in table 7.1 from the most accurate X-ray structure currently available (1M1N [19]). Moreover, (3) I searched the structural data for holes big enough to accommodate one or more water molecules which are also able to transport protons.

$N_{H1}$	Arg
$N_{H2}$	Arg
$N_{\delta 2}$	Asn
$S_{\gamma}$	Cys
$N_{\epsilon 2}$	Gln
$N_Z$	Lys
$O_{\gamma}$	Ser
$O_{\gamma 1}$	Thr
$O_H$	Tyr

**Table 7.1** Amino acids offering mobile proton transport groups and the PDB-label of these atoms.

In order to find paths from the surface of the protein to the cofactor, it had to be decided which atoms belong to the surface. I used an algorithm illustrated in Fig. 7.3: only atoms around the atom of interest  $i$  within a radius of 6 Å were regarded. I summed the normalized vectors of all of these atoms to the atom  $i$ . The resulting vector points away from the sites in the vicinity of atom  $i$ . Thus it points out of the protein if the atom  $i$  belongs to the surface. I regarded a cone of the half opening angle of 45° around this vector and determined if any atom within the 6 Å radius was located within the cone. If this was not the case, then I defined the atom  $i$  to be a surface atom.

In order to find cavities in the protein structure, I scanned it for vacancies. The



**Figure 7.3** Schematic view of the algorithm deciding, which atoms in the PDB database lie on the surface of the protein. Atoms are indicated with black dots. A sphere with the radius of 6 Å is shown as circle. As no atom resides within the 45° cone shown in gray, the central atom is a surface atom. See text for details

scan has been performed on a grid of 0.3 Å. If a sphere with a radius of 3 Å around the grid point is vacant, there is enough space to accommodate a water molecule. This estimation is based on the typical length of a hydrogen bond of 3 Å. Thus an atom in the center of the sphere could form hydrogen bonds to neighboring atoms in the distance of 3 Å.

Indeed I found around 500 of these cavities within the protein tetramer. None of these, however, is located in the direct vicinity of FeMoco.

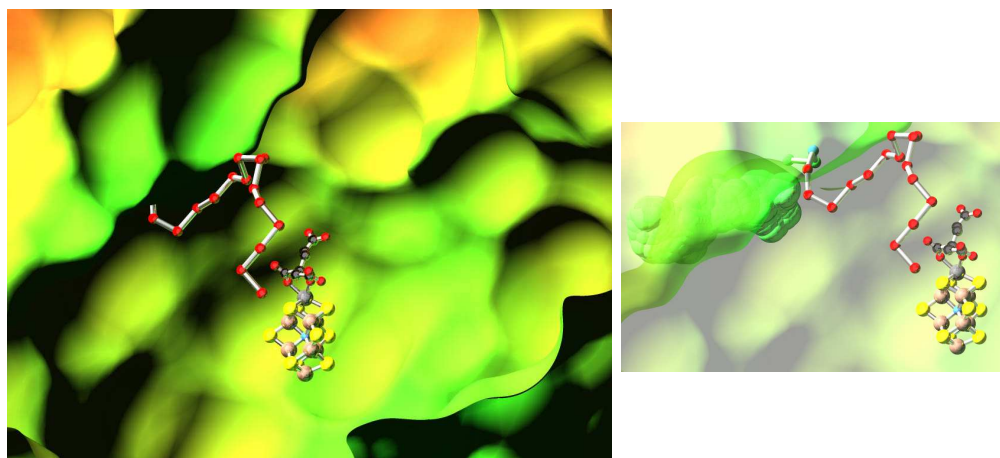
Finally I searched for paths from FeMoco to the surface via atoms, which may take part in the proton transport as described above, in steps of up to 3-3.3 Å.

### 7.1.3 Results

There is only a single proton transfer path from FeMoco to the surface starting from above Fe7 of the FeMoco with the first water molecule being 3 Å to 4.1 Å away from S3B, S4B, and the bridging atom S5A. There are no side branches of the path for the first 14 transport units. Most of these are water molecules except for the protein –OH groups  $O_H$  of Tyr $\alpha$ 446 and  $O_{\gamma 1}$  of Thr $\beta$ 360. At its end, this path is connected to many different branches to reach the surface via 2-5 atoms or cavities. Most of these branches reach the surface via one large cavity between the two halves of the protein tetramer, illustrated in Fig. 7.4.

In the following, all sulfur atoms of the cofactor are discussed in order to show their connections to hydrogen bonds and to the proton transfer path. Hydrogen bond lengths up to 3.3 Å (and 4.0 Å involving S atoms) have been considered. Note, that here only those residues have been considered which can take part in proton transport, as defined above. Therefore there may be hydrogen bonds to all sulfur atoms, but only certain participate in proton channels.

**S3A** No connecting atoms.



**Figure 7.4** The proton path from the surface to the sulfur sites around Fe7. The left figure shows the path, consisting of mostly water atoms (only oxygen shown), the cofactor, and the surface of the protein. In the right figure, the surface is drawn transparent. A cavity between the two halves of the protein tetramer is the starting point of the path. It is shown green.

- S1A** The only chain starting at S1A proceeds via  $S_\gamma$  of Cys $\alpha$ 275 which is incapable of proton transport as it does not exhibit an  $-SH$  group but is bound to Fe1 of FeMoco.
- S4A** No connections to any path, there is only one chain proceeding via Cys $\alpha$ 275.
- S2A** Hydrogen bonds to Arg $\alpha$ 96 and Ser $\alpha$ 278 but those are not connected any further.
- S4B** Is connected to the main path to the surface and will be discussed below.
- S1B** No connecting atoms.
- S2B** No connecting atoms.
- S3B** Shows many connections, among them the main path to the surface.
- S5A** It is connected to a water molecule and to Arg $\alpha$ 96 but these paths end after a few transport units. If the largest considered hydrogen bond distance is increased to 4.1 Å, S5A is connected to the main path leading to the surface.

Extending the criteria for hydrogen-bond distances up to 3.5 Å (and 4.5 Å involving S) does not change the result: S3A, S1A, S4A, S2A, S1B, and S2B still



	atom	residue	distance (Å)
1	S4B		
2	O	HOH 679	3.73
3	O	HOH 15	2.88
4	O	HOH 259	2.84
5	O	HOH 24	2.84
6	O	HOH 50	2.82
7	O <sub>H</sub>	Tyr $\alpha$ 446	2.66
8	O	HOH 26	2.80
9	O	HOH 256	2.71
10	O	HOH 33	2.77
11	O	HOH 222	2.72
12	O <sub><math>\gamma</math>1</sub>	Thr $\beta$ 360	2.79
13	O	HOH 163	2.82
14	O	HOH 438	2.58
15	O	HOH 327	2.89

**Table 7.2** The proton path from the cofactor to the surface. While S4B is given as starting point, the water molecule (679) is also connected to S3B and the bridging sulfur site S5A. HOH denotes water molecules with their respective number in the dataset. The Thr $\beta$ 360 residue belongs to the other half of the protein, the  $\beta_2$  chain.

show only few connections and no path reaching the surface. These extended criteria reveal side branches of the main proton path, which reach the surface in the same area as the main path.

The main path starts from the water molecule with the ordering number 679 in the PDB file, located 3.73 Å from S4B, 4.01 Å from S3B, and 4.02 Å from S5A. The path proceeds via 14 atoms, listed in table 7.2, exhibiting distances of up to 2.89 Å between the oxygen atoms. No vacancies in the protein are included in this path, it consists of ordered water molecules and two amino acid residues. The path is illustrated in Fig. 7.4.

Connected to the last atom listed in table 7.2 there are numerous branches reaching the surface. Some proceed via ordered water molecules, some reach a large pocket between the two halves of the protein. This pocket is connected to the surface.

Thus this path provides the only connection between the surface of the protein and FeMoco, which is capable of transporting a series of protons.

Some atoms without connection to the path, S4A, S2A, and S1A, are located rather close to the protein surface, thus slight conformational changes in the protein may open a path for protons starting from these atoms. With the used definition of surface, S4A is the atom in FeMoco closest to the surface, at a distance of 12.35 Å to the nearest surface atom. However, into that direction, the

protein seems to be very resistant against proton transport: even at hydrogen-bond lengths considered up to 3.9 Å (6.5 Å for sulfur) no paths to the surface can be found.

The path presented here has some transport units in common with the one already present in the literature [65, 68]. However, I extended the proton path from the water pool around homocitrate to sulfur atoms at the cluster. The channel in the literature already ends near the atom number 4 in table 7.2. This, and the atoms 2 and 3 are hydrogen-bound to homocitrate. This extension is crucial since it shows that the proton path is able to provide protons to sulfur atoms and bound substrate near the Fe-sites 7 and 3.

## 7.2 Distance from the cofactor to the surface

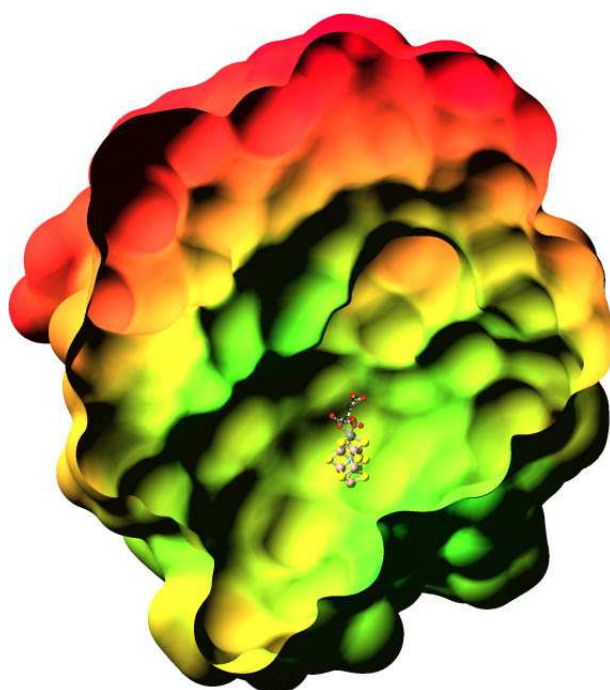
FeMoco is located within the  $\alpha$  subunit of the protein. In order to find possible paths for reactants and products in or out of the protein, the thickness of the protein over the cofactor is of interest. In the following, distances and directions will always be given from the central N atom. While the  $\alpha$ -chain surrounds the cofactor, the  $\beta$ -chain of the MoFe-protein is located approximately into the direction of the Mo atom. The second part of the homo-dimer, i. e. the  $\alpha_2$  and  $\beta_2$  chains, is located approximately into the direction to Fe7. Fig. 7.5 illustrates one half of the protein and the FeMo-cofactor.

I searched for the direction from the nearest surface part to the cofactor. Since the protein surface is all but flat, there are multiple definitions of this surface.

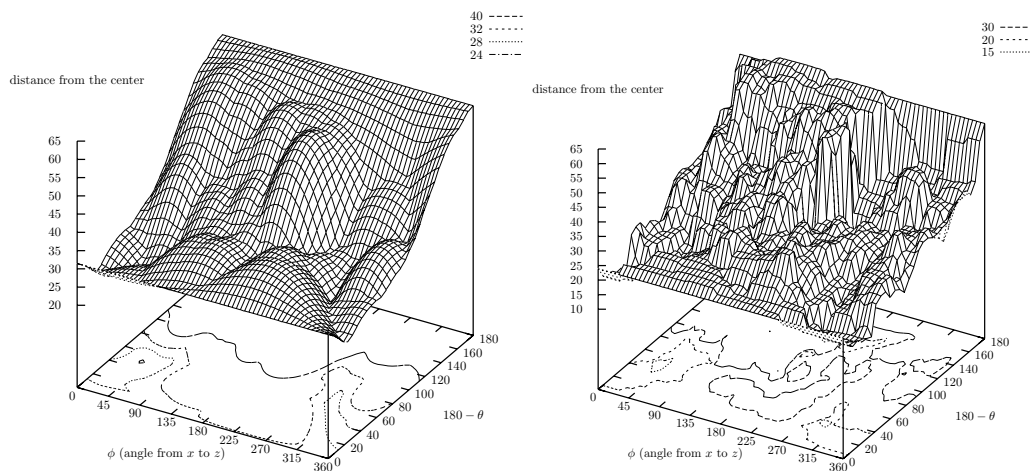
If the distance to the surface of the protein is defined by the distance to a plane going through the outermost atom and perpendicular to the viewing direction, the nearest surface point lies in a direction between Fe2 and S1A and is 24 Å away. Distances defined in this way as function of the polar angles can be seen in Fig. 7.6 on the left side.

A different approach is to regard only atoms within a cylinder of a radius of 4 Å around the viewing direction and find the most distant atom within that. The direction of the nearest surface is similar as above, somewhere in the region between Fe2 and S1A but the distance is only about 13 Å. There is a second “hole” through the protein, a near surface region, into the direction of Fe4. A graph of the thickness of the protein defined by this approach is shown in Fig. 7.6 on the right side.

Thus it can be seen that there is no unique definition of a protein surface. The shortest way from the cofactor to the surface is into the direction from the central ligand to Fe2 or S1A. Reactants like N<sub>2</sub>, however may take this route to the protein surface. This is nearly the opposite direction as the proton transport path.



**Figure 7.5** Illustration of FeMoco and homocitrate in one half (one  $\alpha$  and one  $\beta$  subunit) of the MoFe-protein. The other half connects to the left side. The  $\beta$  surface is cut open. Regions of the surface near the cofactor are indicated green while regions far away from the cofactor are red. The green regions at the lower right are located in the direction from the central ligand to Fe2 and S1A, while the green part at the upper left is a region between the two halves of the protein.



**Figure 7.6** The distance from the center of FeMoco (the central N atom) to the protein surface. In the left graph all atoms are used, i.e. the distance of a plane perpendicular to the viewing direction is given. On the left graph, atoms in a cylinder around the viewing direction are used, see text.  $\theta = 0$  is the direction to the Mo atom,  $\theta = 180^\circ$  is to the Fe1 atom.

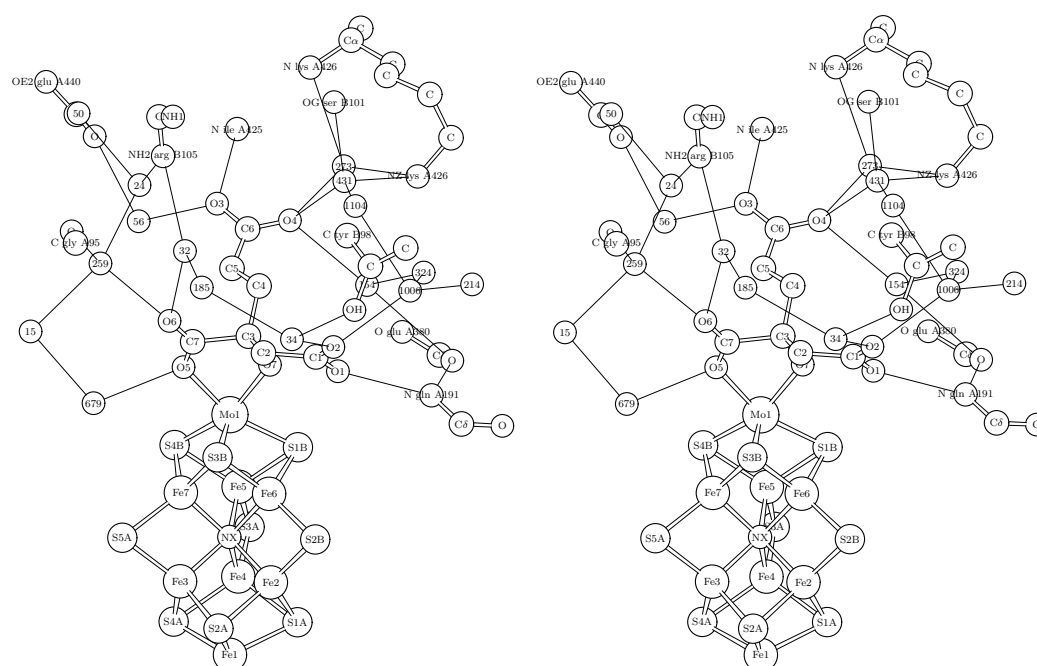
### 7.3 Protonation state of residues

Protons are not contained in the structural databases for nitrogenase as they are difficult to detect by X-ray diffraction experiments. In the DFT calculations, the structure, and thus the protonation state, are important input data. The protonation state of the carboxyl groups of homocitrate is of special importance for possible nitrogen binding to the Mo site of the cofactor. It is possible to obtain some information on the protonation state from bond lengths in the X-ray structure.

The homocitrate, illustrated in Fig. 7.7 allows such an assignment. The functional groups of homocitrate are frequently labeled as “longer carboxyl arm” with the O3 and O4 atoms, the “shorter carboxyl arm” with the O1 and O2 atoms, a carboxyl leg as well as a hydroxy leg.

The distances from the carboxyl C-atom to its oxygen neighbors in the longer carboxyl arm of homocitrate are slightly different. Since the C–O3 distance is longer, it can be concluded that O3 is protonated and that there is a double bond between C and O4. The shorter arm is also most probably protonated, and if, then at O1, the oxygen site pointing towards Fe6. A stereo-view of the hydrogen bond network around homocitrate can be seen in Fig. 7.7.

Arg96 is located near the face of the iron atoms 2, 3, 6, and 7. Its terminal C–N distances are very similar (1.346 and 1.341 Å). Therefore I conclude that both of its terminal nitrogen atoms are protonated. The only hydrogen-bond partner

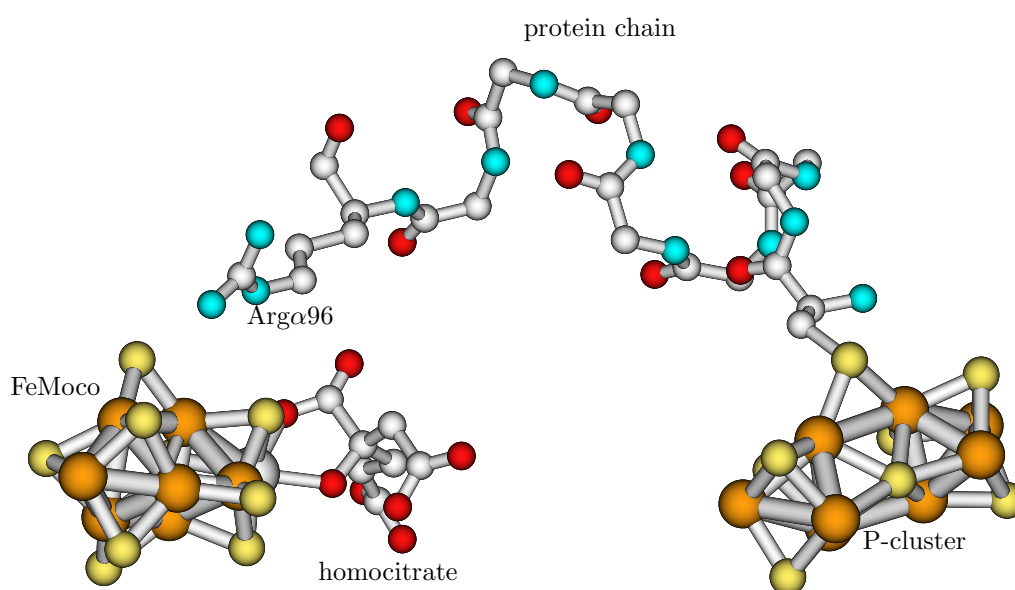


**Figure 7.7** Stereo-view of the network of hydrogen bonds around the homocitrate ligand of FeMoco. Double lines represent covalent bonds while single lines represent hydrogen bonds. In case of water molecules, only their number in the 1M1N data base is given (for the O atom). The water atoms 679, 15, 259, 24, and 50 are part of the proton path.

for one of these, labeled  $N_{H2}$  in the PDB database, is S5A of the cofactor in a distance of 3.276 Å. For the  $N_{\epsilon}$  atom of this residue the only H-bonding partner is S3B, 3.427 Å away. The third nitrogen atom of Arg $\alpha$ 96,  $N_{H1}$ , forms one H-bond to either backbone-O of Gly $\alpha$ 69 or  $O_{\delta 1}$  of Asn $\alpha$ 98.

The side chain of the amino acid arginine Arg $\alpha$ 359 is located centrally above the face spanned by the iron atoms 3, 4, 5, and 7. Although its terminal C-N distances are not completely symmetric, it can be assumed that both N atoms are protonated ( $-NH_2$ ).

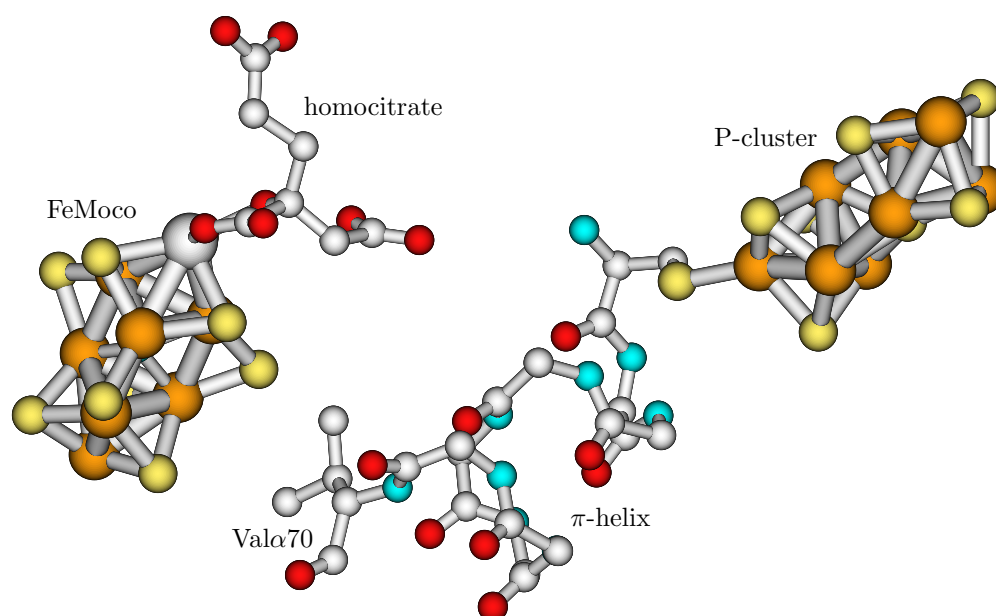
### 7.3.1 Protein chains from the P-cluster to the vicinity of the FeMoco



**Figure 7.8** Arg $\alpha$ 96, above the face spanned by the iron atoms 2, 3, 6, and 7, is connected via a short chain (partially an  $\alpha$ -helix) to the P-cluster via a bridging S atom of Cys $\alpha$ 88. Sequence: Cys Gly Gln Tyr Ser Arg Ala Gly Arg. Only the protein backbone is shown in this picture.

Christiansen [92] mentions that the residues Arg $\alpha$ 96 and Val $\alpha$ 70, which are located in the vicinity of the FeMoco, are both connected via short protein chains to the P-cluster (see Fig. 7.9 and 7.8). Geometric changes in the P-cluster, induced by different oxidation levels, may induce structural changes at these amino acids. These changes may open and close gates to substrate-binding sites at FeMoco, as discussed in section 12.4 on page 173.

Both Arg $\alpha$ 96 and Val $\alpha$ 70 are located above the face spanned by the iron atoms 2, 3, 6, and 7. Rees [10] points out that the helix starting at Arg $\alpha$ 96 is an  $\alpha$ -helix,



**Figure 7.9** Val $\alpha$ 70, above the face spanned by the iron atoms 2, 3, 6, and 7, is connected via a short  $\pi$ -helix to the P-cluster at Cys $\alpha$ 62. Sequence: Cys Ala Tyr Ala Gly Ser Lys Gly Val. Only the protein backbone is shown in this picture.

while the helix starting at Val $\alpha$ 70 is a  $\pi$ -helix, having an extra residue per turn.  $\pi$ -helices are much less common than  $\alpha$ -helices, but frequently occur near active sites and regions that undergo conformational changes [10].

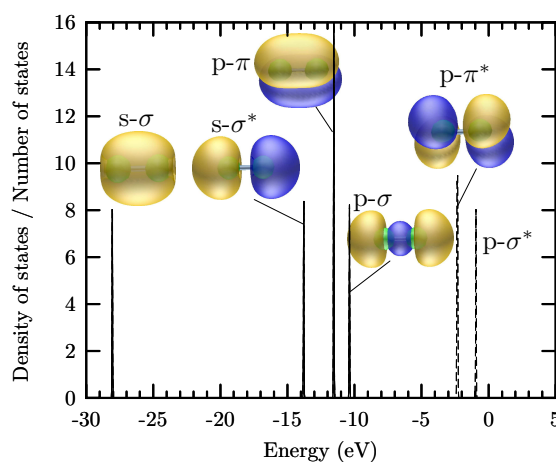


## 8 Dinitrogen and its reduction products in the gas phase

*Everything should be made as simple as possible, but not simpler.*  
Albert Einstein

### 8.1 Electronic structure of dinitrogen

The nitrogen molecule  $\text{N}_2$  exhibits a triple-bond consisting of a  $\sigma$  and two  $\pi$  orbitals which are all occupied. The corresponding antibonding orbitals are empty. The density of states and the shapes of the corresponding orbitals are illustrated in Fig. 8.1.



**Figure 8.1** Density of states and the shape of the corresponding orbitals of the  $\text{N}_2$  molecule. The  $p-\pi^*$  and the  $p-\sigma^*$  states are unoccupied.

In order to weaken the triple bond it is thus possible to reduce or oxidize the molecule. While reduction leads to population of the antibonding  $\pi^*$  orbitals, oxidation de-populated the  $p-\sigma$  or the  $p-\pi$  orbitals.

## 8.2 Comparison to experiment

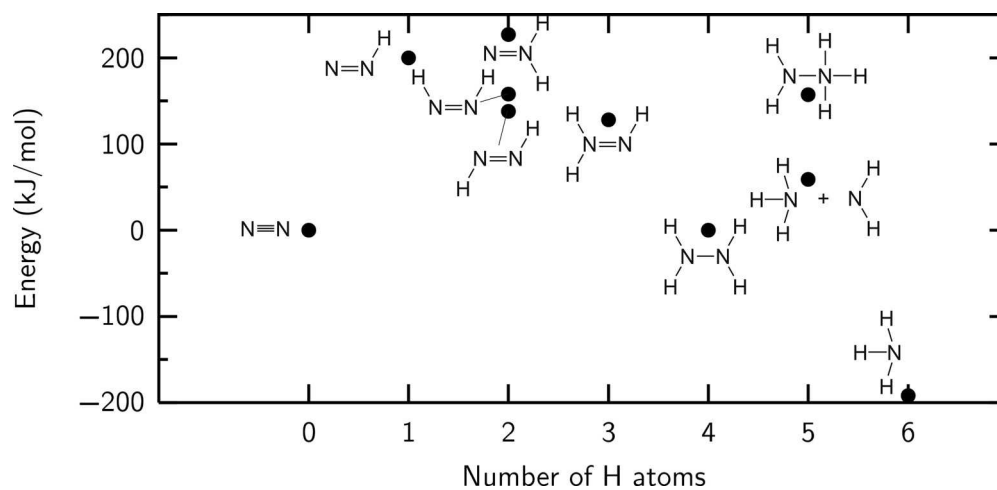
Atomization energies of some of the small molecules involved in dinitrogen reduction are experimentally well established. They can directly be compared to theoretical atomization energies, see table 8.1.

Species	$E_{ZPE}$ [157]	experiment		theory		$\Delta$ PAW–exp.
		$E_{at}$ [120]	$E_{ZPE} + E_{at}$	$E_{at}$ PAW	[121]	
H <sub>2</sub>	24.8	433.0	457.8	444.4	439	–13.4
N <sub>2</sub>	14.7	941.8	956.5	1019.3	1017	62.8
NH <sub>3</sub>	86.7	1157.7	1244.4	1272.4	1264	28.0
NH <sub>2</sub>	48.2	711.3	759.5	798.7		39.2
N <sub>2</sub> H <sub>4</sub>	136.1	1696.2	1832.3	1907.8		75.5

**Table 8.1** Experimental atomization energies  $E_{at}$  (sometimes denoted  $\Sigma D_0$ ) and zero-point vibration energies  $E_{ZPE}$  compared to my calculations (theoretical structure). The sixth column contains theoretical literature data obtained with the PBE functional and experimental geometries. All values in kJ/mol.

Species	spin ( $\hbar$ )	d(N-N) (Å)	E (H)	$E_{rel}$ (kJ/mol)
H	0.5		–0.49875	
H <sub>2</sub>	0		–1.16678	0
NH <sub>2</sub>	0.5		–11.16428	
N	1.5		–9.86257	
N <sub>2</sub>	0	1.105	–20.11336	0
H-N-N	0.5	1.175	–20.62070	200
H-N-N-H <i>trans</i>	0	1.251	–21.22750	138
H-N-N-H <i>cis</i>	0	1.247	–21.21990	158
H <sub>2</sub> -N-N	0	1.214	–21.19374	227
H <sub>2</sub> -N-N-H	0.5	1.353	–21.81459	128
H <sub>2</sub> -N-N-H <sub>2</sub>	0	1.445	–22.44679	0
H <sub>3</sub> -N-N-H <sub>2</sub>	0.5	1.456	–22.97052	157
H <sub>3</sub> -N + N-H <sub>2</sub>	0.5		–23.00774	59
H <sub>3</sub> -N-N-H <sub>3</sub>	0	1.470	–23.48679	333
N-H <sub>3</sub>	0		–11.84346	–192

**Table 8.2** Calculated data concerning N<sub>2</sub> and its reduction products. Relative energies are given with respect to N<sub>2</sub> and  $x\text{H}_2$ .



**Figure 8.2** Calculated data concerning Relative energies  $\text{N}_2$  and its reduction products with respect to  $\text{N}_2$  and  $x\text{H}_2$ .



## 9 The resting state

*If it looks like a duck, and quacks like a duck, we have at least to consider the possibility that we have a small aquatic bird of the family anatidae on our hands.*

Douglas Adams

### 9.1 Determination of the charge state

Experiment provides the atomic structure of the FeMo cofactor. Additionally to that, however, the charge state has to be known for a detailed understanding and for reliable calculations of the reaction mechanism. Here I will show that the charge state of the resting state is given by  $[\text{MoFe}_7\text{S}_9\text{N}]^{\pm 0}$ .

The charge state will always be referred as a “core charge” of the  $\text{MoFe}_7\text{S}_9\text{N}$  subunit. A charge of  $-3e$  is assigned to the ligands: SH ( $-1e$ ) imidazole (neutral) and glycolate ( $-2e$ ). Thus the charge of the whole cell in the calculation is the core charge minus three  $e$ .

Usual determinations of any physical properties are guided by the variational principle of the total energy. Its minimum is searched. If this scheme is directly applied to the cofactor, the charge state it would have in vacuum would be calculated. The cofactor, however, is surrounded by the protein. On the one hand this acts like a dielectric, stabilizing charges, and on the other hand it is partially charged itself. There are basic amino acids, like for example histidine, around the cofactor which are positively charged and acid amino acids which are negatively charged at neutral pH 7. Even the local pH value in the direct vicinity of the cofactor is not known. All these facts change the electronic chemical potential,  $\mu_e = \partial E / \partial n$  with  $n$  being the number of electrons, around the cluster from that of the vacuum. Thus the state with the lowest energy in vacuum is not necessarily the ground state in the protein.

Physical properties other than the total energy have to be used to find the charge state. The total spin and the structure, as well as hyperfine parameters of specific sites, may be compared to experiment. I used the former two for identifying the resting state.

### 9.1.1 Computational details

I performed DFT [113, 115] calculations based on the projector augmented wave [122, 124] (PAW) method as described in section 5.2. The implementation into the CP-PAW program, developed by my advisor, Peter E. Blöchl, has been used. I used the gradient-corrected PBE [121] functional for exchange and correlation.

The cofactor has been modeled by truncating its ligands as shown in Fig. 9.5 on page 107 for example. Cysteine has been replaced by SH, histidine by imidazole protonated at its uncoordinated nitrogen atom, and homocitrate by glycolate. Thus no double bonds or aromatic systems have been broken. The first two coordination shells of Mo are identical to those in the protein. No atoms of the cofactor core are protonated, which is rationalized in section 10.2.1 on page 125.

The numbers of pairs of partial waves ( $\phi_i, \tilde{\phi}_i$ ) and projector functions ( $\tilde{p}_i$ ) used for each angular momentum ( $\ell, m$ ) are given in table 9.1. Convergence of this setup has been tested.

Element	core	s	p	d
Fe	[Ar]	2	2	2
Mo	[Kr]	3	3	2
S	[Ne]	2	2	2
O	[He]	2	2	1
N	[He]	2	2	1
C	[He]	2	2	1
H	none	2	1	0

**Table 9.1** Core configuration and number of projector functions and pairs of partial waves per angular-momentum state ( $\ell, m$ ) with s,p and d-character as used throughout the work for this thesis.

A plane-wave cutoff of 30 Ry for the wave functions and of 60 Ry for the density have been used.

The plane-wave based PAW method leads to the occurrence of periodic images of the structures. The electrostatic interaction between them has explicitly been subtracted [158]. Wave function overlap has been avoided by choosing a unit cell large enough to keep a distance of more than 6 Å between periodic images.

Atomic structures have been optimized by damped Car-Parrinello [130] molecular dynamics and all degrees of freedom have been relaxed. The convergence has been tested by monitoring if the kinetic temperature remains below 5 K during a simulation of 50 fs (200 time steps). During that simulation no friction has been applied to the atomic motion and the friction on the wave function dynamics has been chosen sufficiently low to avoid a noticeable effect on the atomic motion. In most cases, the kinetic temperature remained below 1 K during this final test.

All calculations allowed for non-collinear spin densities, except especially noted.

I used a collinear description to investigate the 10 different spin orderings BS1 to BS10 shown in Fig. 9.6. I achieved the desired spin ordering by applying an artificial potential to the wave function. It has only been applied in the first 30 steps of the calculation and removed afterwards. This potential acts with reversed sign at the two spin directions. It is localized within the augmentation region of the selected atoms. Its energy  $E_U$  is given by

$$E_U = \sum_{i,j} D_{i,j} \langle \phi_i | U_0 \mathcal{P}_\ell e^{-(r/r_c)^2} | \phi_j \rangle \quad (9.1)$$

with

$$D_{i,j} = \sum_n f_n \langle \tilde{\Psi}_n | \tilde{p}_i \rangle \langle \tilde{p}_j | \tilde{\Psi}_n \rangle, \quad (9.2)$$

$$\mathcal{P}_\ell = \sum_m \delta(|\vec{r}| - |\vec{r}'|) Y_{\ell,m}(\vec{r}) Y_{\ell,m}(\vec{r}'). \quad (9.3)$$

Here,  $\phi_i$  and  $\phi_j$  are one-center partial waves of the respective atom. All 7 iron atoms have been used.  $U_0$  is a constant defining the strength of the potential. For all atoms to be forced to a spin-up configuration, I chose  $U_0$  to be  $-1.5 \text{ H}$  for spin-up wave functions and  $+1.5 \text{ H}$  for spin-down wave functions. The sign has been reversed, if a spin-down atom is should be achieved. As the potential is only applied to the d-part of the wave function,  $\ell = 2$  has been chosen in  $\mathcal{P}_\ell$ . I used a cutoff-radius  $r_c$  of  $2 a_B$ .  $f_n$  denotes the occupations of state  $n$ .  $\tilde{\Psi}_n$  is the auxiliary wave function,  $\tilde{p}_i$  and  $\tilde{p}_j$  are the projector functions, explained in section 5.2.2.  $Y_{\ell,m}$  denote spherical harmonics.

The nature of chemical bonds can be analyzed based on the electronic structure. The concept is similar to the COOP (Crystal Orbital Overlap Population) analysis of R. Hoffmann [159]. What I refer as COOP here, is the energy resolved density matrix elements projected onto individual atomic orbitals. Negative COOP indicates bonding contribution to the corresponding bond, positive COOP indicates antibonding contribution. The COOP between two orbitals  $\chi$  and  $\chi'$  is defined as

$$D_{\chi,\chi'}(\epsilon) = \sum_n \langle \chi | \Psi_n \rangle \delta(\epsilon - \epsilon_n) \langle \Psi_n | \chi' \rangle. \quad (9.4)$$

The energy is denoted as  $\epsilon$ ,  $\Psi_n$  is the Kohn-Sham wavefunction of state  $n$ .  $\epsilon_n$  is the energy of that state. The  $\chi$  and  $\chi'$  are atomic-like orbitals. I used the one-center partial waves  $\phi_i$ , truncated at the atomic sphere radius. Equation (9.4) defines the atom-projected density of states if  $\chi = \chi'$ .

### 9.1.2 Total spin as criterion for finding the resting state

An experimental total spin of  $S = 3/2$  of the resting state, as described in section 2.2.1, implies an odd number of electrons within the cluster. Thus all even charge state are candidates for the resting state.

Charge states with an odd electron number ranging from  $-2e$  to  $+4e$  have been investigated. Assuming oxidation states of  $\text{Mo}^{4+}$ ,  $\text{S}^{2-}$ , and  $\text{N}^{3-}$  these correspond to one to seven ferric iron sites ( $\text{Fe}^{3+}$ ), the other ones being ferrous ( $\text{Fe}^{2+}$ ). Note that the assignment of integer oxidation numbers to individual atoms is somewhat arbitrary and no direct result of calculations, as the electrons are distributed over the whole cluster. Thus all reasonable charge states have been investigated.

Among the charge states with an odd electron number ranging from  $-2e$  to  $+4e$ , I found that only the charge state  $0e$ , which is collinear, can clearly be identified with an  $S = 3/2$  spin state. In this state, three iron sites are in their ferric state, while the remaining four are ferrous. Charges of  $+2e$  as well as  $+4e$  result in an  $S = 1/2$  state, and the charge state of  $-2e$  has a non-collinear spin distribution with  $S = 0.24$ . Non-collinear spin states generally result in a total spin differing from half-integer numbers as described in section 5.5.4. Structural relaxation coupled to relaxation of the spin ordering is important for obtaining reliable information on the total spin as the spin ordering strongly depends on the atomic structures.

charge ( $e$ )	spin ( $\hbar$ )	
-2	0.229	non-collinear
-1	1.07	non-collinear
$\pm 0$	$3/2$	
+1	0.591	non-collinear
+2	$1/2$	
+3	0	
+4	$1/2$	
+5	0	

**Table 9.2** Total spin of different charge states of FeMoco

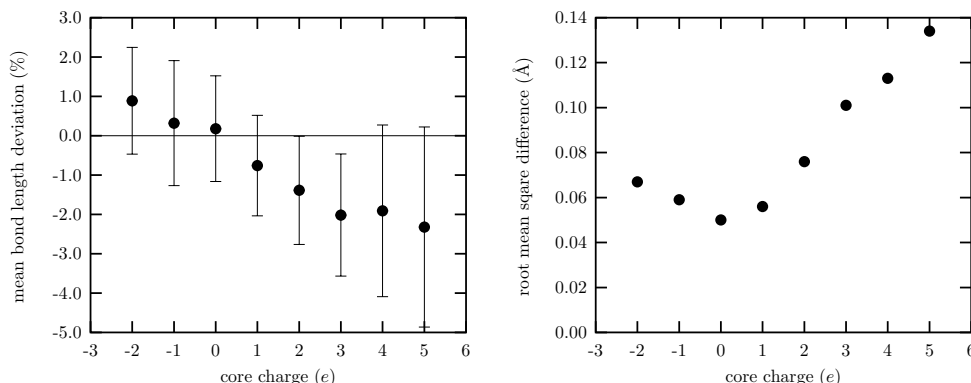
Thus the charge states  $+2$  and  $+4$  can clearly be excluded as candidates for the resting state. No clear decision can be made for the most reduced state  $-2$ . The DFT result of  $S = 0.24$  may correspond to a physically meaningful state of  $S = 1/2$ , but  $S = 3/2$  cannot clearly be excluded.

Thus, concerning the spin state, a core charge of zero most probably refers to the resting state.

### 9.1.3 Atomic structure as criterion for finding the resting state

The atomic structure of the isolated cofactor generally expands upon reduction and contracts upon oxidation. This fact is illustrated in Fig. 9.1 as well as in Fig. 9.2. In both parts of Fig. 9.1, all atomic distances within the  $\text{MoFe}_7\text{S}_9\text{N}$  unit are considered. The distances obtained from theory are compared to those from





**Figure 9.1** Change of the cluster size at different charges. Left: the mean fractional deviation of calculated distances in various oxidation states relative to the X-ray structure [19]. All distances in  $\text{MoFe}_7\text{S}_9\text{N}$  are considered. The error bars indicate the root mean square deviation from the mean value. Right: root mean square difference in bond length to the X-ray structure.

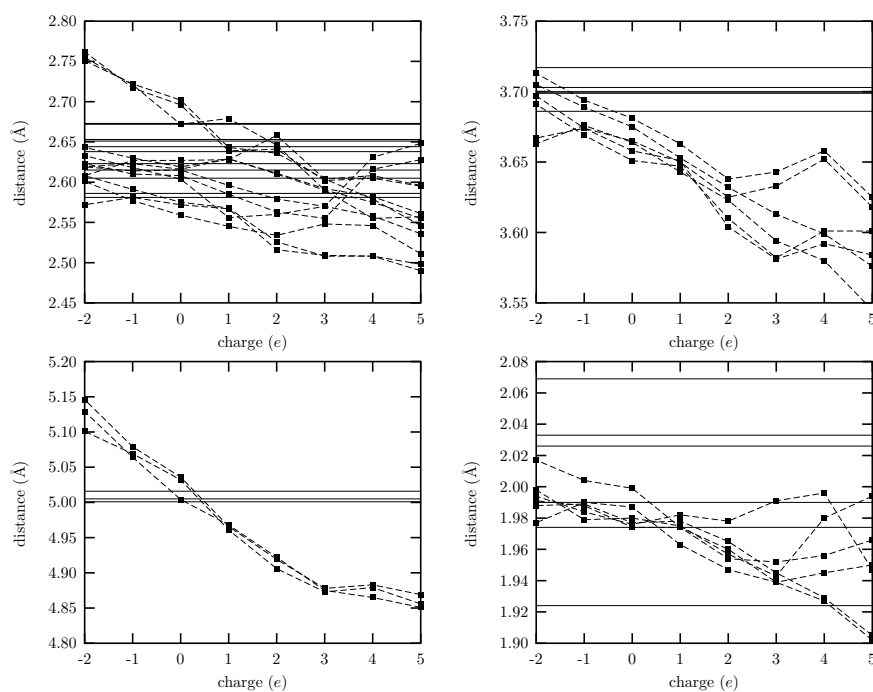
crystallographic analysis, namely the first entry in the 1M1N PDB structure [19]. The fractional deviations are averaged over all 153 of these distances and plotted in the left graph. The error bars correspond to the root mean square deviation of individual distances from this mean value. In the right panel the root mean square difference of bond length and longer distances from the experimental values is directly given.

Both parts of Fig. 9.1 not only show the contraction upon oxidation but also show that a charge of zero fits best to the experimental structure. The structural difference in percent is lowest at this charge as well as the root mean square difference. The large error bars at higher charge states in the left graph also indicate a distortion of the structure. In these cases, the central cage of the cofactor loses its symmetry. Bonds from the central ligand to some iron sites contract.

The structure of the neutral charge state perfectly matches the experimental structure as it can be seen from their overlay in Fig. 9.3 on page 105. Only rotation of the truncated ligands are different between these structures. These are truncation effects.

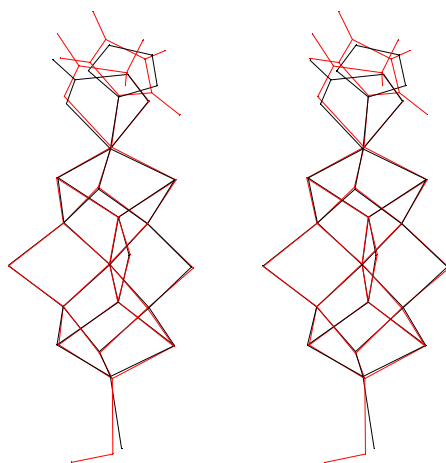
The conclusion of analysis of the total spin as well as the atomic structure is the assignment of a neutral core charge to the resting state:  $[\text{MoFe}_7\text{S}_9\text{N}]^0$ .

My finding of a neutrally charged resting state is in agreement with other theoretical work. Comparison of the mean Mössbauer isomer shifts [25] as well as reduction potentials [24, 25] with experiment result in the same assignment. Hinemann and Nørskov [46] used a charge state with the same formal oxidation



**Figure 9.2** The Fe–Fe bond length at different charge states. The first graph shows all short bonds of directly bond irons. The second one shows the diagonals in the Fe-prism and the third one the distances from  $\text{Fe}_1$  to  $\text{Fe}_{5-7}$ . Broken lines are calculated results, solid lines the X-ray distances and filled squares the collinear state at charge  $-3$ . The last graph shows the Fe–N bond lengths.

states of the Fe-sites. However, they considered all three  $\mu_2$ -sulfur bridges of the cofactor to be protonated.

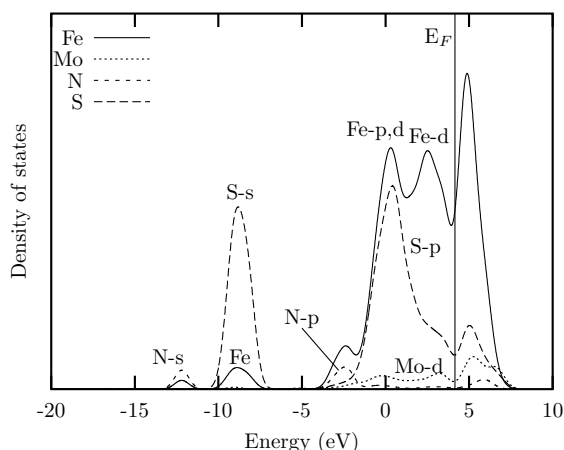


**Figure 9.3** Theoretical vs. experimental structure: stereoview of the superimposed structures from X-ray diffraction (black) and the theoretical resting state (red).

## 9.2 Analysis of the resting state

### 9.2.1 Electronic and spin structure

An overview of the contributions of the individual atomic species to the electronic density of states is given in Fig. 9.4. The main chemical bonds determining the cluster structure can be seen. Before discussing them, see section 9.2.3, I will concentrate on the spin structure and use it to determine the formal oxidation states.

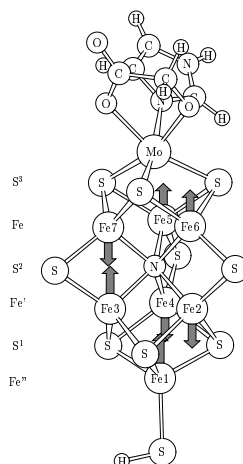


**Figure 9.4** Overview of the electronic density of states of FeMoco in its resting state projected onto the S, Fe, Mo atoms and the central ligand (N).  $\sigma$ -bonds between S and Fe can be seen around  $-9$  eV,  $\pi$ -backdonation around  $0$  eV. The Fe-S antibonding states are out of range of this graph.

The seven iron sites of FeMoco are in their high-spin state, resulting in a high spin-polarization in the density of states, see Fig. 9.7 on page 109. This means that each Fe-site exhibits four or five unpaired electrons with parallel spin depending on its oxidation state. The local spin on the other atoms can be neglected with the possible exception of Mo, discussed below. Fig. 9.5 shows the spin ordering of the ground state.

In contrast to collinear calculations, a non-collinear description of the spin density, as applied in all my calculations except explicitly noted, generally allows to find the ground state of this complex system in one run.

The fact that the spin ordering shown in Fig. 9.5 is indeed the ground state can be verified by calculating the energy of the other possible spin orderings. Such an analysis has been performed by Lovell et al. [72] for the vacant cofactor. They assumed a three-fold rotational symmetry of the cluster and labeled the 10 possible spin orderings with BS1 to BS10, explained in Fig. 9.6. In order to result in an

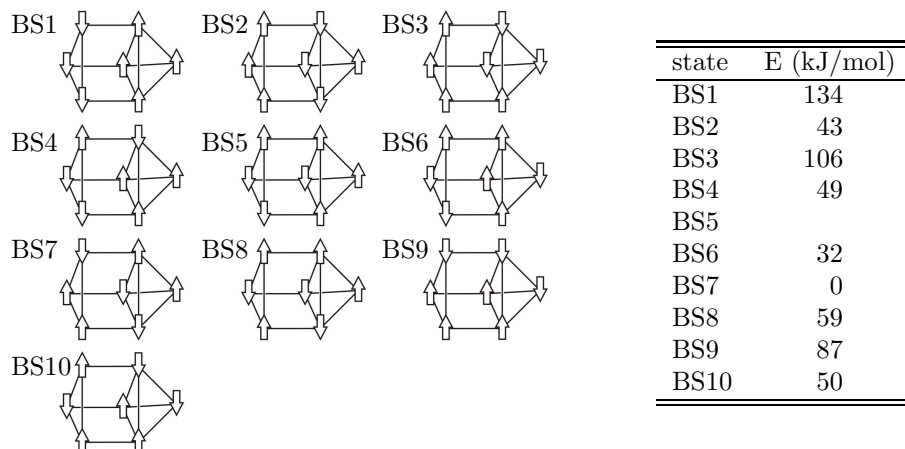


**Figure 9.5** Spin ordering of the resting state.

$S = 3/2$  system, four sites have to be spin-up and three spin-down. Lovell et al. calculated the energies of the cofactor without central ligand in these spin states. They concluded that state BS2 had the lowest energy but state BS6, with the second-lowest energy was more consistent with the experimental structural data. BS6 has subsequently been used in a number of theoretical studies [25, 46, 104] on the cluster with central ligand.

In the structure with the central ligand, however, the spin ordering with the lowest energy is BS7. I calculated all 10 of these states with collinear spin description. All cases but BS5, which converged to BS7, converged to the desired spin ordering. This shows that large barriers for spin-flips exist in collinear calculations. The energies of the different states are given in Fig. 9.6 on the following page. The ground state and the first excited state BS6 are well separated by an energy difference of 32 kJ/mol. It is important to note that all energies have been obtained by relaxing the structure. The energetic order of the states may be different without structural relaxations. Different spin orderings lead to remarkable differences in the atomic structures.

For the state with the lowest energy, BS7, I also calculated the resulting three states removing the three-fold rotational symmetry of the cluster. Their energies are given in table 9.3. They differ by less than 7.1 kJ/mol. In my structural model, the three-fold rotational symmetry is only broken by the ligands of the Mo site. Thus the energetic differences are much smaller than the expected differences induced by the protein environment.



**Figure 9.6** All 10 spin orderings of FeMoco consistent with a total spin of  $S = 3/2$ . A three-fold rotational symmetry has been assumed for the definition of these 10 states. Breaking of this symmetry will be discussed in the text. The energies of the respective states relative to BS7 are given.

state	E (exp <sup>1</sup> )	E
BS7	(kJ/mol)	(kJ/mol)
Fe <sub>5</sub>	0.7	4.0
Fe <sub>6</sub>	3.4	7.1
Fe <sub>7</sub>	0	0

<sup>1</sup> at the experimental structure

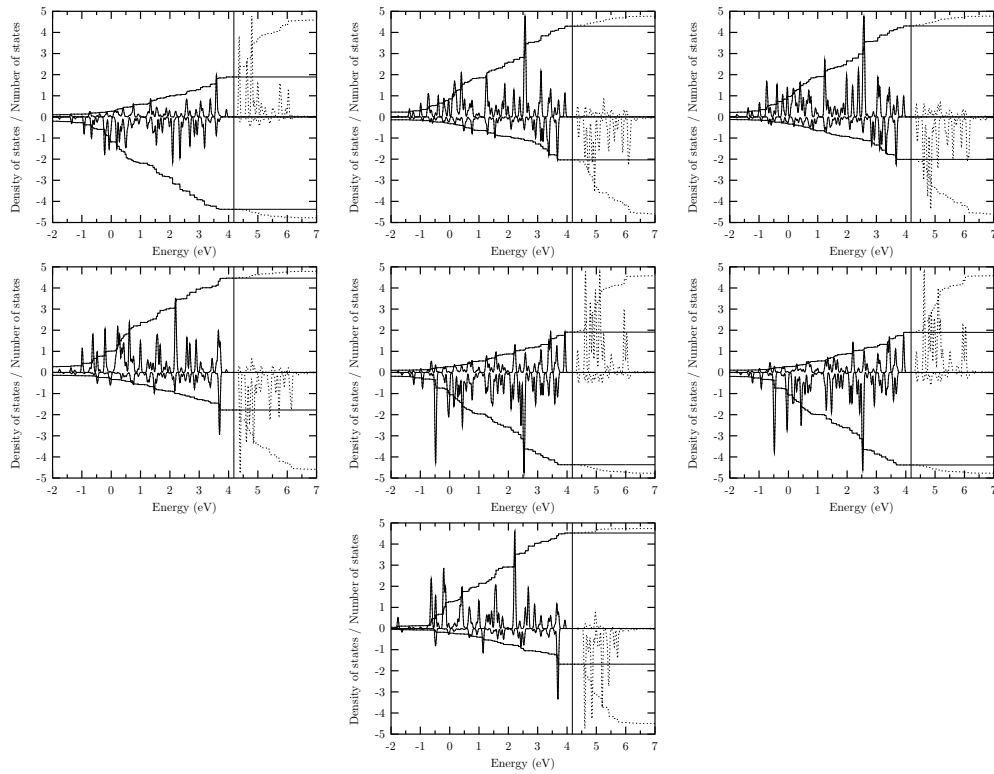
**Table 9.3** Energies of the three sub-states of BS7, differing by rotation around the molecule's axis. The number of the unpaired spin-down atom is given. Prior to structural relaxation, the energetic differences are even smaller.

### Assignment of formal oxidation numbers to the metal atoms

Although the net charge of the cluster would suggest formal oxidation numbers like  $[\text{Mo}^{4+}(\text{Fe}^{3+})_3(\text{Fe}^{2+})_4\text{S}_9\text{N}^{3+}]^0$  a detailed analysis of the non-bonding metal orbitals leads to an assignment of  $[\text{Mo}^{4.5+}(\text{Fe}^{2.5+})_5(\text{Fe}^{2+})_2\text{S}_9\text{N}^{3+}]^0$  with the two ferric sites being Fe<sub>5</sub> and Fe<sub>6</sub>.

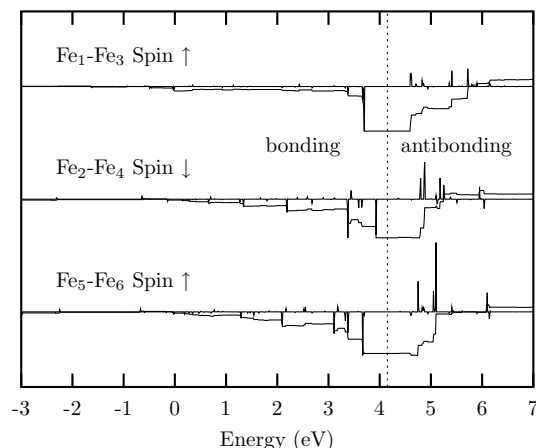
This assignment can be rationalized by investigating the metal-metal bonds formed by electrons not involved in the metal-ligand bonding.

A ferrous Fe-site and a ferric Fe-site with parallel spins can form a mixed-valence bond as described in sections 2.2.3 and 6.1. Both majority-spin d-orbitals are filled and the remaining minority-spin electron is distributed between the iron sites, forming the bond. This bond is weak because it is formed by only one electron.



**Figure 9.7** Projected densities of states in the resting state: iron sites. On the top the DOS projected on d-states of Fe<sub>7</sub>, Fe<sub>5</sub> and Fe<sub>6</sub>. Then Fe<sub>3</sub>, Fe<sub>4</sub> and Fe<sub>2</sub> and finally at the bottom Fe<sub>1</sub>. The ordering is the same as in the structure in Fig. 9.5. Fe<sub>7</sub> is the only atom not forming a mixed-valence pair with another iron site.

Three pairs of iron sites with parallel spins can be found in the spin structure of the resting state as it can be seen in Fig. 9.5. The occurrence of a bond between them can be confirmed by calculating the COOPs between them, similar to the overlap between their atomic orbitals. Bonding and antibonding contributions of each one-particle state can be seen in a COOP plot.



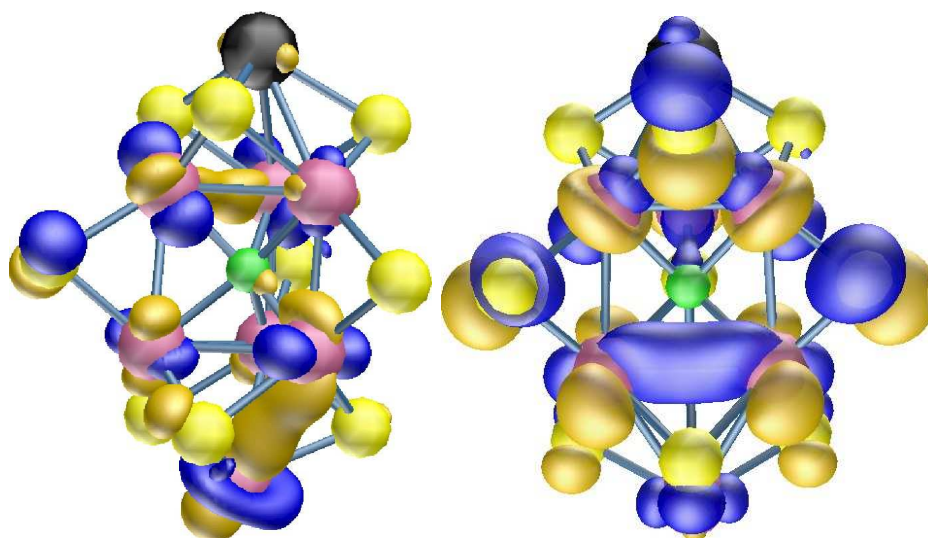
**Figure 9.8** Fe-Fe bonds: COOPs of the minority spin  $\sigma$ -bonds between ferromagnetically coupled iron pairs. The negative peaks indicate bonding states while the positive peaks indicate antibonding states. The dotted line at 4.15 eV indicates the Fermi energy.

Such plots are shown for the iron-iron bonds in Fig. 9.8. The bonding states lie below the Fermi energy and thus are occupied while the antibonding states are unoccupied. Thus clear binding contributions can be found in the minority spin direction of each parallel iron pair. It can be seen from the COOPs that many of the states below the Fermi level contribute to the binding. This is because the electrons are not localized in one bond but distributed over the whole cluster. Fig. 9.9 shows isosurfaces of two one-particle states with significant contributions to these metal-metal bonds.

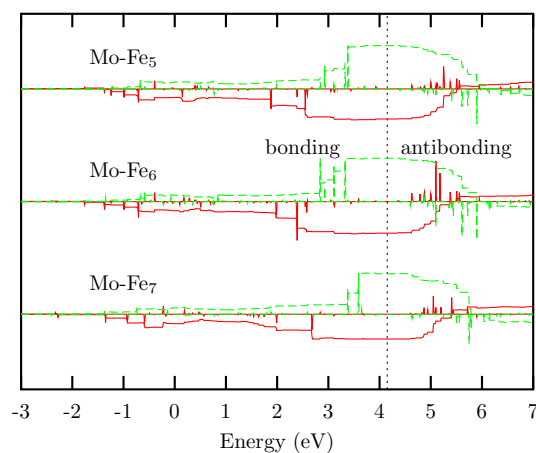
Now that the bonding network between the iron atoms is known, I focus on the molybdenum site. The left picture of Fig. 9.11 shows a clear bonding contribution between the molybdenum atom and the minority-spin d-orbital of site Fe<sub>7</sub>. This site does not take part in the metal-metal bond network of the other Fe-sites. The Mo orbital participating in this bond is a d-orbital with  $t_{2g}$  symmetry in the approximate octahedral environment of the Mo site, thus a low-lying orbital.

The other two  $t_{2g}$  orbitals of Mo are also partially occupied. It can be seen from the COOP graphs in Fig. 9.10 that there are also metal-metal bonds between the Mo site and its other two iron neighbors, Fe<sub>5</sub> and Fe<sub>6</sub>. This can also be seen in the iso-surface of the total spin density of the cluster shown in Fig. 9.11 on the

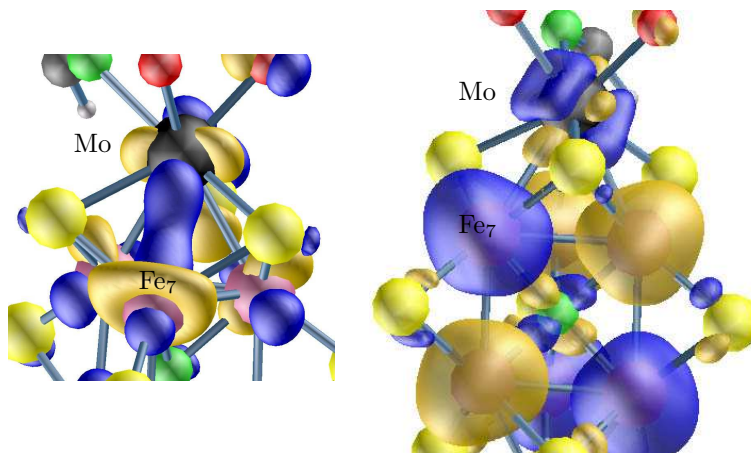




**Figure 9.9** Metal-metal bonds between iron atoms. Iso-surfaces of the one-particle states which contribute most to the iron-iron bonds are shown. The left picture shows a high-lying occupied spin-down state, covering the bonds between  $\text{Fe}_1$  and  $\text{Fe}_3$  as well as  $\text{Fe}_6$  and  $\text{Fe}_7$ . The right picture shows the highest occupied spin-up state, the bond between  $\text{Fe}_2$  and  $\text{Fe}_4$ . It can be seen from the COOP graphs in Fig. 9.8 that these are not the only states contributing to the respective bonds.



**Figure 9.10** Mo-Fe bonds: COOPs of both spin directions are shown. The majority direction of the respective iron site is shown in full lines (red), the minority direction in dashed lines (green).



**Figure 9.11** Left: a spin-up state shows the metal-metal bond between d-electrons of  $\text{Fe}_7$  and Mo. The total spin density, right, not only shows some spin density at the central ligand but also at the Mo site. All three of its  $t_{2g}$  orbitals are partially occupied, one with spin-up (yellow) and two with spin-down (blue) electrons. These electrons form the metal-metal bonds between Mo and Fe.

right side. One spin-up d-orbital (yellow) forms the bond to the spin-down  $\text{Fe}_7$  while two spin-down d-orbitals (blue) connect to  $\text{Fe}_5$  and  $\text{Fe}_6$ .

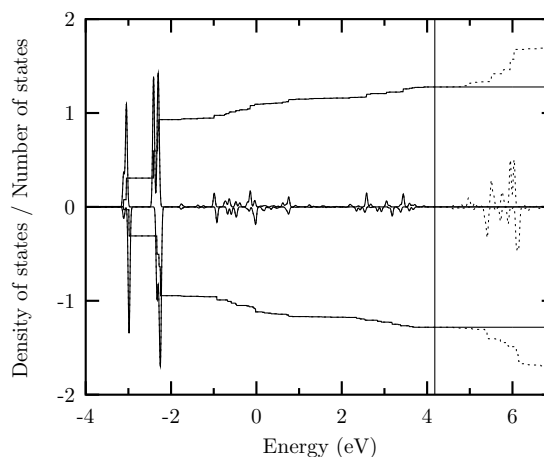
In order to obtain formal oxidation states, I use the following thought experiment: I start with all partially filled d-shells empty, thus  $\text{Mo}^{6+}(\text{Fe}^{3+})_7$ . Electrons of a metal-metal bond are equally distributed between the metal atoms. Thus Mo, taking part in three bonds to all its iron neighbors, gets three half electrons resulting in  $\text{Mo}^{4.5+}$ . The iron sites  $\text{Fe}_5$  and  $\text{Fe}_6$  exhibit a bond between them and a bond to Mo from each site, thus they obtain one electron leading to  $\text{Fe}^{2+}$  for both of them. All other iron sites take part in one bond and thus are formally assigned  $\text{Fe}^{2.5+}$ .

As the spin direction of each metal-metal bond is fixed to the minority direction of the participating atoms, the number of electrons have to sum up to the correct total spin. The assignment given above is consistent with both the total spin and the total charge.

The formal distribution of one electron over two atoms results in a spin of half an electron on the Mo site. This is consistent with the DFT calculations, which result in a small spin-density at Mo.

Although the given assignment is rationalized based on the total spin density and on one-particle states, other assignments may be reasonable. The ferrous character of sites  $\text{Fe}_5$  and  $\text{Fe}_6$  may be distributed to the other spin-up sites  $\text{Fe}_1$  and  $\text{Fe}_3$ . On the other hand the Mo atom may also be described as being in its diamagnetic 4+ state with  $\text{Fe}_7$  being ferric. However, the given assignment results

in similar oxidation states for all iron sites which explains the similar experimental Mössbauer parameters given in section 2.2.1 on page 11 and compared to my results in section 9.2.5.



**Figure 9.12** Density of states projected on the p-states of the central nitrogen ligand.

The central nitrogen atom may safely be regarded as being in the  $3-$  state. In the density of states (Fig. 9.12), there are some unoccupied p-states. The origin of these lies in the complex-bonds to the iron sites. They are empty antibonding orbitals. In a formal representation one would attribute the electrons of these bonds to nitrogen, which results in an  $N^{3-}$  ion.

### 9.2.2 Atomic structure

The atomic structure perfectly matches the experimental structure for the cofactor core as it can be seen from Fig. 9.3 and table 9.4. The bond lengths of Mo to the histidine and homocitrate ligands differ by 4.3% from experiment. This is caused by the weakness of these bonds and possible forces of the environment acting on the ligands.

### 9.2.3 Chemical bonds

While the spin structure and the formal oxidation states of the metal sites are determined by metal-metal bonds, metal-ligand bonds are responsible for the atomic structure of the cofactor.

The Fe sites are in a distorted tetrahedral environment. This effects the spin structure: tetrahedral coordinated iron site are known experimentally to exhibit a high-spin state [160]. All ligands of the terminal Fe site are sulfur atoms, while

	experiment				theory		
	X-ray diffr.			EXAFS	min	max	mean
	2.0 Å	1.6 Å	1.16 Å	[37]			
Mo-O	2.02	2.32	2.182	2.142	2.047	2.198	2.12 (−2.8%)
Mo-N	2.13	2.48	2.304	2.142	2.403	2.403	2.40 ( 4.3%)
Mo-S	2.24	2.34	2.345	2.342	2.369	2.410	2.39 ( 1.7%)
Mo-Fe	2.63	2.69	2.696	2.719	2.695	2.780	2.74 ( 1.5%)
Mo-Fe'		5.06	5.062	5.060	5.093	5.118	5.10 ( 0.8%)
Mo-Fe''	6.89	6.97	7.004		7.123	7.123	7.12 ( 1.7%)
Fe-S <sup>3</sup>	2.32	2.24	2.237	2.225	2.248	2.267	2.26 ( 0.8%)
Fe-S <sup>2</sup>	2.46	2.21	2.208	2.225	2.196	2.204	2.20 (−0.4%)
Fe'-S <sup>2</sup>	2.46	2.23	2.221	2.225	2.200	2.208	2.20 (−0.7%)
Fe'-S <sup>1</sup>	2.35	2.26	2.269	2.225	2.236	2.305	2.28 ( 0.4%)
Fe''-S <sup>1</sup>		2.28	2.268	2.225	2.220	2.310	2.28 ( 0.5%)
Fe-Fe	2.52	2.65	2.622	2.612	2.581	2.612	2.60 (−0.8%)
Fe'-Fe'	2.59	2.68	2.657	2.612	2.622	2.645	2.63 (−1.0%)
Fe-Fe'	2.55	2.61	2.594	2.612	2.556	2.602	2.59 (−0.3%)
Fe-Fe' d	3.60	3.73	3.700	3.660	3.661	3.691	3.68 (−0.6%)
Fe-N			2.003		1.966	1.999	1.99 (−0.8%)

**Table 9.4** The structure of the resting state of FeMoco compared to experiments. Exploiting the approximate three-fold rotational symmetry of the cluster, averages of corresponding bond length are given. The labeling of the site refers to Fig. 9.5. Additionally to the mean value of the theoretical bond length, the minimum and maximum of the corresponding bond length as well as its deviation from the most accurate X-ray structure [19] is given. The X-ray structures refer to [11], [23], and [19], respectively.

one ligand of the prismatic sites is the central nitrogen ligand. The bonds between metals and non-metals show  $\sigma$  as well as  $\pi$  character. Formally, both  $\sigma$  electrons are provided by the ligand, as this is a complex-bond. The  $\pi$  electrons partially occupy the sulfur-d states and are provided by the metal. Therefore the  $\pi$  part of the bond mainly exists in the minority spin direction of the participating atom. This kind of bonding is usually referred as  $\sigma$ -donor  $\pi$ -acceptor bonding or back-donation.

The bonds between the iron sites and the central ligand are rather weak and show nearly exclusively  $\sigma$  character. The p-states of the central nitrogen site are separated into a  $p_z$  state pointing into the rotation axes of FeMoco and two states,  $p_x$  and  $p_y$ , 0.7 eV higher in energy as  $p_z$ , as it can be seen in Fig. 9.12 on page 113.

The molybdenum site shows a slightly distorted octahedral environment with three sulfur, one nitrogen, and two oxygen ligands. The  $t_{2g}$  orbitals point to the neighboring iron sites. As they are energetically low, they are partially occupied and are responsible for the metal-metal bonds to the iron sites.

#### 9.2.4 Which atom is in the center?

In the whole discussion up to now and in all calculations, nitrogen has been assumed as being the central ligand although the results from crystallographic analysis left it open to be N, C, or O. While nitrogen can be rationalized as central ligand because it is present at the conversion process, it cannot be excluded from experiment that a different species is the central ligand and is not exchanged in the catalytic process.

The atomic structure suites best for comparing these three candidates for the central ligand in a calculation. Keeping the sum oxidation state of all the other atoms fixed, I tried  $C^{4-}$ ,  $N^{3-}$ , and  $O^{2-}$  as central ligands. The results are listed in table 9.5: while the bond length of the cluster with C and N compare well with experiment, oxygen generally expands the structure, most conspicuous at the iron-iron distances. Therefore oxygen can be ruled out as candidate for being the central ligand. No answer can be given concerning carbon or nitrogen.

It can be assumed that the chemistry of the cluster would not significantly be changed by replacing nitrogen with carbon as they have approximately the same size and are isoelectronic.

experiment		theory		
[19]		C	N	O
Mo-O	2.182	2.15 (−1.7%)	2.12 (−2.8%)	
Mo-N	2.304	2.37 ( 2.8%)	2.40 ( 4.3%)	
Mo-S	2.345	2.38 ( 1.4%)	2.39 ( 1.7%)	2.40 ( 2.3%)
Mo-Fe	2.696	2.72 ( 0.9%)	2.74 ( 1.5%)	2.74 ( 1.6%)
Mo-Fe'	5.062	5.09 ( 0.6%)	5.10 ( 0.8%)	5.21 ( 2.9%)
Mo-Fe''	7.004	7.11 ( 1.5%)	7.12 ( 1.7%)	7.05 ( 0.7%)
Fe-S <sup>3</sup>	2.237	2.24 ( 0.1%)	2.26 ( 0.8%)	2.24 ( 0.2%)
Fe-S <sup>2</sup>	2.208	2.19 (−0.7%)	2.20 (−0.4%)	2.21 ( 0.0%)
Fe'-S <sup>2</sup>	2.221	2.21 (−0.7%)	2.20 (−0.7%)	2.20 (−1.1%)
Fe'-S <sup>1</sup>	2.269	2.28 ( 0.6%)	2.28 ( 0.4%)	2.28 ( 0.6%)
Fe''-S <sup>1</sup>	2.268	2.28 ( 0.6%)	2.28 ( 0.5%)	2.25 (−1.0%)
Fe-Fe	2.622	2.58 (−1.5%)	2.60 (−0.8%)	2.71 ( 3.3%)
Fe'-Fe'	2.657	2.64 (−0.5%)	2.63 (−1.0%)	2.84 ( 6.9%)
Fe-Fe'	2.594	2.58 (−0.5%)	2.59 (−0.3%)	2.69 ( 3.9%)
Fe-Fe' d	3.700	3.67 (−0.7%)	3.68 (−0.6%)	3.86 ( 4.4%)
Fe-X <sup>1</sup>	2.003	1.99 (−0.9%)	1.99 (−0.8%)	2.09 ( 4.5%)

<sup>1</sup>X=C, N, O

**Table 9.5** Bond length with C, N, or O as central ligand compared to the X-ray structure from the 1M1N data base.

### 9.2.5 Hyperfine parameters

Mössbauer isomer shifts  $\delta$  and electric field gradients  $\Delta E_Q$  have been calculated for the iron sites.

The electric field gradient (EFG) can be calculated from the electron density [161]. It is a traceless tensor  $V_{ij}$ . After diagonalization the three eigenvalues are sorted as  $|V_{zz}| \geq |V_{yy}| \geq |V_{xx}|$ . Hence,  $V_{zz}$  is the eigenvalue of the EFG tensor with the largest absolute value. The asymmetry parameter  $\eta$  is defined by the remaining two eigenvalues  $V_{xx}$ ,  $V_{yy}$ :

$$\eta = \frac{|V_{yy}| - |V_{xx}|}{|V_{zz}|} \quad (9.5)$$

The electric field gradient interacts with the quadrupole moment of the iron atoms. This interaction can be measured by Mössbauer spectroscopy as quadrupole splitting ( $\Delta E_Q$ ) which is proportional to  $V_{zz}$  [161]. For conversion, I used a quadrupole moment of  $^{57}\text{Fe}$  of  $Q = 0.16 \text{ b}^1$  [162]. As the electric field gradient depends stronger on the accuracy of the calculation than the total energy, it has been converged by increasing the cutoff for the plane waves to 40 Ry and for the density to 160 Ry.

Note that in case of high asymmetry ( $\eta$ ) the sign of the quadrupole splitting  $\Delta E_Q$  is not well-defined any more. For  $\eta = 1$  the smallest eigenvalue of the EFG tensor has the same absolute value as the largest eigenvalue, because the eigenvalues sum up to zero. In this case small changes in the electric field gradient may change the sign of  $\Delta E_Q$ .

theory	$\delta$	$\Delta E_Q$	$\eta$	exp	$\delta$	$\Delta E_Q$	$\eta$
Fe1	0.46	0.90	0.75	A <sup>1</sup>	0.39	-0.69	1
Fe2	0.34	0.67	0.92	A <sup>2</sup>	0.48	-0.94	1
Fe3	0.37	1.48	0.95	A <sup>3</sup>	0.39	-0.56	1
Fe4	0.35	0.66	0.88	A <sup>4</sup>	0.41	0.68	1
Fe5	0.36	-0.95	0.57	B <sup>1</sup>	0.33	-0.66	0.9
Fe6	0.35	-0.89	0.45	B <sup>1</sup>	0.33	-0.66	0.9
Fe7	0.30	-1.36	0.78	B <sup>2</sup>	0.50	-0.65	1

**Table 9.6** Mössbauer isomer shifts  $\delta$  and electric field gradients  $\Delta E_Q$  of the iron sites. Theoretical results are compared to experiment [30].  $\delta$  and  $\Delta E_Q$  are given in mm/s.

In table 9.6 the results of the calculations are compared to experiment. A site-per-site assignment of the hyperfine data is not possible because of the similarity of the iron sites between each other.

---

<sup>1</sup>1 b=10<sup>-28</sup> m<sup>2</sup>

Five sites show the limiting case of  $\eta = 1$  in experiment. The calculations also show a high asymmetry, the mean value is 0.76, but still smaller than the experimental mean value of 0.97. This is an acceptable agreement.

I compare the absolute values of the quadrupole splitting  $\Delta E_Q$  because of the above-mentioned uncertainties of its sign for high asymmetry. The calculated quadrupole splitting is with an average of 0.99 mm/s a little larger than the experimental mean value of 0.69 mm/s. Also the standard deviation of the theoretical results is with 0.32 mm/s larger than the experimental standard deviation of 0.12 mm/s. This is a good agreement as errors in the range of 50% are not uncommon in EFG calculations.

The isomer shift  $\delta$  is a linear function of the electron density at the nucleus. As no accurate ab-initio conversion function is available, I have calculated  $\delta$  by calibrating it to simple iron-sulfur clusters with known charges and known isomer shifts. The theoretical mean value is with 0.36 close to the experimental one with 0.40. In both cases, there is one site (Fe1, B<sup>2</sup>) with remarkably higher isomer shift than the average. In the calculation, this site is the terminal iron atom. Its chemical environment is different from that of the other Fe-sites which is represented in the isomer shift.



## 10 Protonation

*When a philosopher says something that is true then it is trivial.  
When he says something that is not trivial then it is false.*

Carl Friedrich Gauss

Now that the charge state of the resting state has been determined, the first steps in the catalytic cycle of biological nitrogen fixation, protonations of the cofactor, are to be investigated. In this chapter it will be shown that protons are usually only added to the sulfur sites bridging two iron sites. I will show that the resting state is not protonated and will rationalize the experimentally observed contraction upon reduction of the cofactor. The experimentally observed  $\text{H}_2$  production, on the one hand in the absence of substrate and on the other hand as a side reaction of  $\text{N}_2$  fixation, will be explained mechanistically.

### 10.1 Protonation sites

The FeMoco cluster may, in principle, be protonated at any electronegative atom. I investigated these possibilities as well as protonations at metal atoms. In the latter case one would formally speak of hydride ( $\text{H}^-$ ) formation. Protonation at the Mo site is no stable minimum.

In the following, individual protonation sites will be discussed in the order of decreasing stability. In these calculations, one proton and one electron have been added to the resting state. The reduction will be rationalized in section 10.2.3. Protonation energies are given in table 10.1, the corresponding structures are shown in Fig. 10.1 on the following page.

**S<sup>2</sup>:** Protonation of the sulfur sites bridging two iron atoms (labeled S5A, S2B, and S3A in the 1M1N data base [19]) is energetically most favorable. The sulfur site is not coordinated trigonally-planar but the S–H bond is nearly perpendicular to the Fe–S bonds. This lowest-energy structure is collinear with  $S = 2$ .

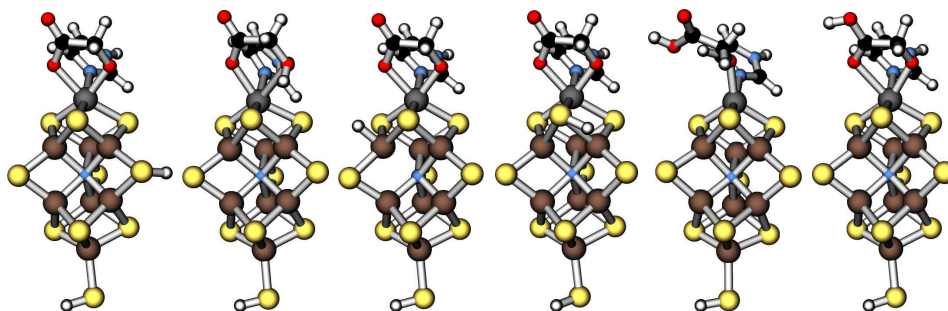
**O<sub>7</sub>:** Protonation of the alcoholate O of homocitrate leads to a significantly enlargement of its bond to Mo. However, the bond is not broken: the bond length increases from 2.06 Å to 2.60 Å.

**Fe:** Addition of a proton to one of the prismatic Fe sites formally results in an oxidation of this site and hydride formation. It pulls the iron atom somewhat out of the cage. However, the  $N_x$ -Fe bond is not completely broken. The iron site stays trigonal-bipyramidally coordinated. Hydride formation can easily lead to  $H_2$  production as it will be shown below.

**S<sup>3</sup>:** Protonation of a sulfur site bridging three iron atoms weakens the S-Fe bond. The proton is positioned somehow between S and Fe. As the energy decreases, it can be assumed that the proton is transferred to an iron site where it forms a hydride.

**O<sub>5</sub>:** Protonation of the carboxyl O bound to Mo leads to breaking of the Mo-O bond. The homocitrate becomes monodentate<sup>1</sup>.

**O<sub>6</sub>:** Protonation of the carboxyl O of homocitrate not coordinated to Mo increases the Mo-O<sub>5</sub> distance from 2.23 Å to 2.49 Å. However, the energy surface of this bond is very flat, stronger enlargement of the bond or even its cleavage cannot be ruled out. The spin ordering is non-collinear. Consequences of protonation of homocitrate oxygen atoms on a possible hydrogen bond from homocitrate to His $\alpha$ 442 will be discussed in section 10.1.1.



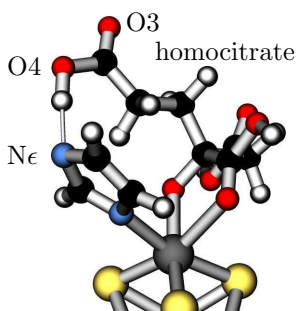
**Figure 10.1** Different protonation sites of FeMoco. From left to right: the lowest energy structure  $S^2$ ,  $O_7$ , Fe,  $S^3$ ,  $O_5$ , and  $O_6$ .

### 10.1.1 Protonation and monodentate homocitrate

Protonation of the carboxyl O of homocitrate which is not connected to the molybdenum ( $O_6$ ) has been widely discussed in the literature [67, 69, 99]. Even though it can be seen from table 10.1 that this protonation has the highest energy of all sites near the cofactor, I will discuss it and compare the results to mechanisms

site	E (kJ/mol)	$S$ ( $\hbar$ )
S <sup>2</sup>	0	2
O <sub>7</sub>	11	0.90
Fe	19	0.19
S <sup>3</sup>	47	2
O <sub>5</sub>	65	0.26
O <sub>6</sub>	102	0.18

**Table 10.1** First protonation energies at the cofactor after one reduction step at different protonation sites.



**Figure 10.2** Possible hydrogen bond between homocitrate and histidine.

proposed in the literature.

Grönberg et al. [99] described the possibility of forming a hydrogen bond between N $\epsilon$  of His $\alpha$ 442 and O4 of the longer carboxyl arm of homocitrate as shown in Fig. 10.2. As a consequence it is expected that homocitrate loses one bond to Mo and becomes monodentate. The freed coordination site at Mo is a prerequisite for N<sub>2</sub> coordination to Mo.

I investigated this system with collinear calculations keeping the spin state fixed: spin transitions are not expected to occur during hydrogen bond formations. For this study I expanded my model to include the whole homocitrate. To define the starting geometry for these calculations, I used the 1M1N PDB data base entry [19]. The geometry has been fully relaxed except for those atoms, at which the amino acids have been truncated (S $\gamma$  of Cys $\alpha$ 275 and C $\gamma$  of His $\alpha$ 442). The latter have been frozen in space.

The X-ray data only specify the positions of the heavy atoms. Based on the elongated C–O distances in the X-ray data, I assumed that the homocitrate is protonated at O3 of the longer carboxyl arm and at O1 of shorter carboxyl arm as discussed in section 7.3.

According to my calculations, the transition from bidentate to monodentate

<sup>1</sup>Monodentate means that homocitrate offers only a single coordination site to Mo. In contrast to that, it offers two coordination sites in the bidentate mode.

homocitrate requires both protonation of O6 and a hydrogen bond between O4 and N $\epsilon$  of histidine. This hydrogen bond pulls homocitrate into the monodentate coordination. With an intact hydrogen bond, the transition from bidentate to monodentate homocitrate is approximately iso-energetic. However, while I can stabilize the hydrogen bond by placing the proton at O4, the smaller acidity of histidine suggests that the proton resides at N $\epsilon$  of histidine. In that case, however, the hydrogen bond breaks, so that homocitrate remains bidentate.

As a side remark, this finding depends on the presence of the central ligand. In its absence, homocitrate is already stable in its monodentate coordination, when only O6 is protonated.

### 10.1.2 Protonation and H<sub>2</sub> production

FeMoco may produce H<sub>2</sub> after at least two reduction and protonation steps with respect to the resting state. This reaction is exothermic by at least 84 kJ/mol. This is the case when both protons are bound to S<sup>2</sup> sites. In case of other protonation sites, it is even stronger exothermic. This H<sub>2</sub> production competes with N<sub>2</sub> fixation as it takes two reduction equivalents from the cofactor per H<sub>2</sub> evolved and leaves it back in the resting state. However, a cofactor in a highly reduced and protonated state is required for effective N<sub>2</sub> binding as it will be discussed in section 11.1.5.

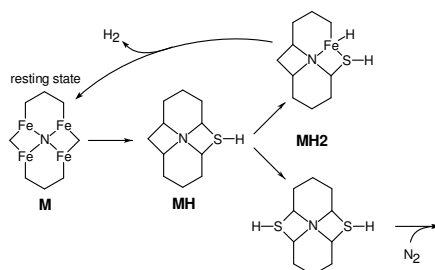
On the one hand, the weak binding of protons to FeMoco must be regarded as an important ingredient to the ability of the cofactor to reduce N<sub>2</sub>. If protons were strongly bound, they would not be transferred from the proton path to N<sub>2</sub> but to the cofactor itself.

On the other hand, the destability of the protonated cofactor with respect to H<sub>2</sub> production seems to be inconsistent with experiment. The Thorneley-Lowe scheme, obtained from measurements of the reaction kinetics, states that at least two protons have to be bound to the cofactor prior to N<sub>2</sub> binding.

Here, steric effects play an important role. In the case mentioned above, with two protons at different S<sup>2</sup> sites, the distance between the protons is at least 5.4 Å. This is too much for a direct interaction between them. A proton shuttle mechanism is also unlikely to occur, as one proton with positive partial charge would have to be transferred to another one with positive partial charge. Water molecules would be required for the proton shuttle mechanism. However, there are no water molecules in the protein structure between the two S<sup>2</sup> sites accessible by proton channels, S5A and S2B, see section 7.1.3. Therefore, H<sub>2</sub> production from two protonated S<sup>2</sup> sites is kinetically suppressed. It can also be expected that the system with all three S<sup>2</sup> sites protonated will be kinetically stabilized even though it is thermodynamically unstable. However, this situation is unlikely as there are proton channels to only two sites.

In contrast to that, hydride formation in an aqueous system usually leads to

H<sub>2</sub> production. Whenever two protons are delivered by the same channel, the first one binds to the respective S<sup>2</sup> site, but the second one forms a hydride at the nearest Fe site. Double protonation of one S<sup>2</sup> bridge is not stable. The proton and the neighboring hydride may easily form H<sub>2</sub>, a reaction exothermic by 89 kJ/mol with a barrier of 35 kJ/mol.



**Figure 10.3** Mechanism for the first two reduction and protonation steps of FeMoco. The cofactor is drawn schematically by including the central ligand, iron sites, and the sulfur bridges between them. Starting from the resting state *M*, the first proton is added to an S<sup>2</sup> sulfur bridge (*MH*). If the second one is delivered by the same channel as the first one, hydride is formed and H<sub>2</sub> evolves. Otherwise it remains stable and the cofactor is ready for N<sub>2</sub> binding.

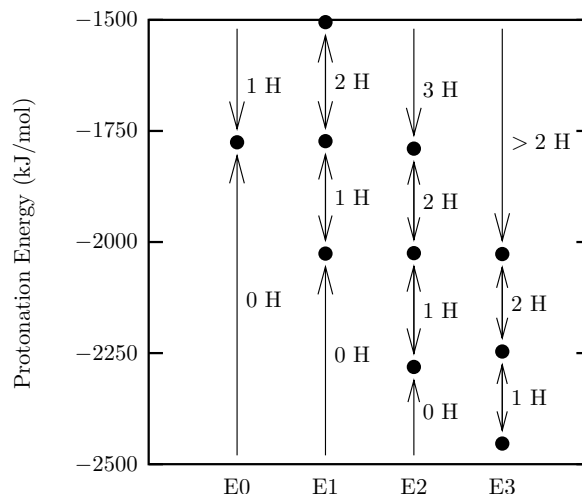
This reaction, double proton delivery by one channel, is avoided by the enzyme. S5A, one of the two S<sup>2</sup> sites accessible by proton channels, has two positively charged histidine residues in the direct vicinity. These reduce the basicity of the site, as pointed out by Durrant [68]. Therefore S2B, the other S<sup>2</sup> site, next to the channel which can only transfer one proton, is protonated first. After that, atom S5A will be protonated. If no suitable substrate is available at that stage, the next proton forms a hydride and produces H<sub>2</sub>.

This mechanism explains two experimental observations:

1. Ordinary hydrogen evolution competing with N<sub>2</sub> reduction as discussed in section 3.2.1 on page 20. The mechanism is consistent with non-stoichiometric production of H<sub>2</sub>.
2. H<sub>2</sub> is produced whenever no other substrate is available.

## 10.2 The chemical potential of protons

Now that the protonation sites are known to be the sulfur atoms bridging two iron sites, in the following I will discuss the sequence of reduction and protonation steps. In this discussion I will not consider protonation at the homocitrate ligand.



**Figure 10.4** Protonation energies for different charge states of the cofactor.

While in principle the protonations of FeMoco are independent of the reduction state, the protonation energy changes with the charge state as shown in Fig. 10.4. In this figure, E0 represents the resting state, E1, E2, ... are states reduced by the respective number of electrons. It can clearly be seen that the protonation energies increase by approximately 250 kJ/mol per proton added to the cofactor and decrease by the same amount for each electron added.

The results shown in Fig. 10.4 have been obtained from collinear calculations as the spin state is not expected to significantly change on protonation. For each reduction state (each  $E_x$ ) the spin state has been optimized and kept for the different protonation states, as they have the same number of electrons. Note that here, different charge states of the isolated molecule are compared. In the real system, the protein, energies given in the figure will be affected by electrostatic screening by the environment. This means that the energy differences may be scaled down by a factor of about 3–5, the assumed dielectric constant of a protein. Each row of dots in the figure represents one total charge state, thus the qualitative result of the figure remains unchanged.

Protonation transfer occurs much faster than electron transfer, which limits the rate of the complete reaction. Thus, for each reduction state, an equilibrium concerning the number of protonations will be reached. The number of transferred protons depends on the chemical potential of protons  $\mu_p$ . Only, if  $\mu_p$  is larger than the protonation energy, the proton will be transferred.

The chemical potential of protons does not vary in space nor in time as it is defined as an equilibrium quantity. If the proton supply through the path is slower than the consumption by the cofactor, there is no equilibrium and thus no

well-defined  $\mu_p$ .

It can be seen from Fig. 10.4 that any value for  $\mu_p$  results in a ping-pong mechanism for proton and electron addition to the cluster. Each reduction step induces a single proton transfer to the cofactor.

In the following, I will estimate  $\mu_p$  by relating theoretical and experimental results.

### 10.2.1 Protonation state of the resting state

	E0H0	E0H1	X-ray [19]
d(S–Fe <sub>7</sub> )	2.196	2.296	2.207
d(S–Fe <sub>3</sub> )	2.208	2.273	2.214
∠(Fe–S–Fe)	70.95	66.91	71.59
∠(Fe <sub>3</sub> –Fe <sub>7</sub> )	2.556	2.518	2.586

**Table 10.2** Comparison between the unprotonated resting state, the protonated resting state, and the X-ray structure. Distances are given in Å and angles in degrees.

The question, whether the resting state is protonated, can be addressed by comparing the calculated atomic geometries with experiment as given in table 10.2. The whole cluster contracts by about 0.5 % upon protonation of one sulfur bridge under preservation of the charge state. This contraction is mainly driven by a decrease of the Fe–S–Fe angle of the protonated sulfur bridge. It also results in smaller Fe–Fe distances. As the agreement of the unprotonated cluster structure with X-ray [19] and EXAFS [34, 37] experiments is deteriorated by protonation, I conclude that the resting state is unprotonated. This fact provides, when combined with the results depicted in Fig. 10.4, an upper bound for the proton chemical potential in the cavity at the protonation energy of the resting state, around  $-1750$  kJ/mol.

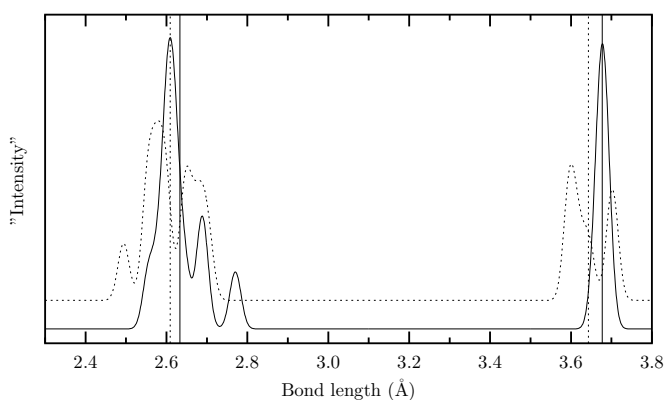
### 10.2.2 Protonation state of the first reduced state

The protonation state of the first reduced state can also be determined by comparison of the structural changes to experiment. EXAFS measurements indicate that the cluster contracts upon reduction by one electron in *Azotobacter vinelandii* (Av) [34], while no significant changes have been found for *Klebsiella pneumoniae* (Kp) [33].

While electron transfer alone does not change the structure of the cofactor appreciably in my calculations, protonation decreases the angle of the sulfur bridges, which in turn contract the cluster. Metal-metal distances, as measurable by EXAFS, are given in table 10.3 for the resting state and for the reduced states E1H0

	resting state		reduced state			EXAFS
	EXAFS	E0H0	E1H0	E1H1		
				H(S <sup>2</sup> )	H(O <sub>7</sub> )	
d(Fe–Fe)	2.63	2.633	2.636	2.609	2.628	2.60
d(Fe–Fe)	3.74	3.677	3.679	3.643	3.662	3.72
d(Mo–Fe)	2.71	2.736	2.742	2.713	2.710	2.65

**Table 10.3** Change of bond length upon reduction, measured by EXAFS [34] in *Azotobacter vinelandii* and compared to my DFT calculations. All length are given in Å. It can be seen that protonation of the reduced cluster is required to explain the experimentally observed contraction.



**Figure 10.5** Distribution of the Fe–Fe bond length of the resting state (full line) and the protonated reduced state (E1H1, dashed line, shifted upwards). Each occurring length is broadened in order to simulate a measurement of the “length distribution” in EXAFS. Mean values are given as vertical lines. The differences of the mean values represent the contraction of the cage upon reduction.

and E1H1. Although the changes are small, it can clearly be seen that the cluster contracts upon protonation on S<sup>2</sup> as well as O<sub>7</sub>, the alcoholate of the homocitrate ligand. The contraction may be explained by the fact the the proton withdraws negative charge from the cluster. This partial oxidation contracts the cluster. The structural changes only weakly depend on the protonation site. The reduction of the mean bond length upon protonation is also depicted in Fig. 10.5 where each occurring distance has been broadened to mimic a measured intensity. It can be seen that the EXAFS measurements in *Azotobacter vinelandii* clearly indicate a protonation after the first reduction step.<sup>2</sup>

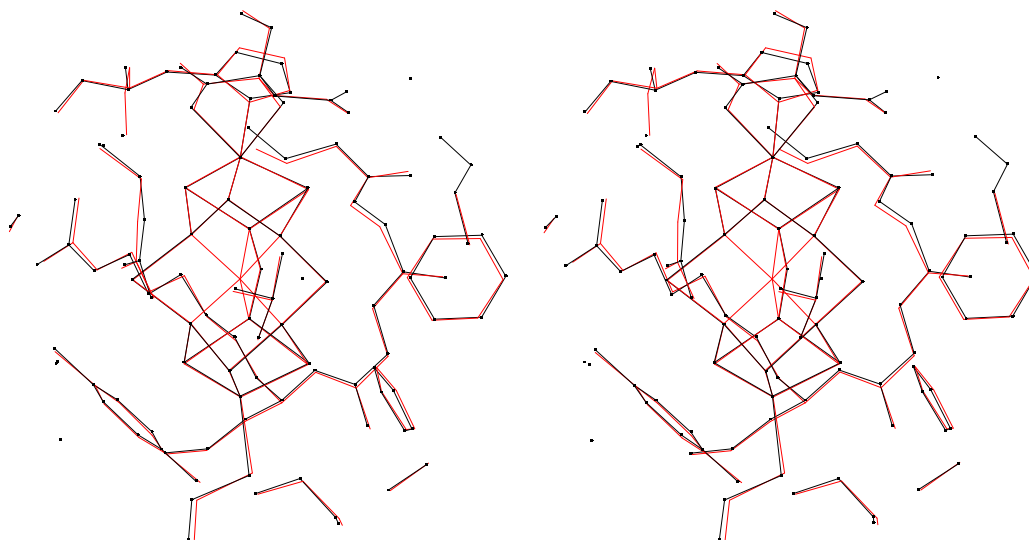
<sup>2</sup>The fact that this really was the first reduction step has been verified by EPR measurements carried out simultaneously with the EXAFS measurements [34].



The fact that no significant structural changes can be measured in *Klebsiella pneumoniae* [33] indicates that no protonation takes place in this species. Proton transfer, as apparent by the contraction, depends on subtle changes of the protein between different bacteria's nitrogenases. As both have the same functionality, the question if the first state is already protonated must be quite on the border for the two species, slightly on the protonated side for *Azotobacter vinelandii* and slightly on the non-protonated side for *Klebsiella pneumoniae*. This allows to identify the proton chemical potential approximately with the first protonation energy of the cofactor reduced by one electron.

Comparison of the protein environments of these two species, see Fig. 10.6, does not provide a clear hint for rationalizing the difference in protonation behavior. There are some minor difference in the distances of the bridging S atoms to protein residues. Differences bigger than 0.05 Å in distance are: (1) S2B to C of Gln $\alpha$ 191 (which cannot be a proton source) is 4.637 Å in Av and 4.576 Å in Kp; (2) S5A to N of Arg $\alpha$ 96 (which is positively charged) is 3.276 Å in Av and 3.224 Å in Kp, and (3) S5A to a water molecule is 3.440 Å in Av and 3.377 Å in Kp.

The most relevant changes occur near the S5A and N of Arg $\alpha$ 96 as well as the water molecule are closer to S5A in Kp. Close distance to an arginine nitrogen can reduce the hydrogen affinity due to the positive charge of arginine while close distance to water is assumed to be important for protonation. From the first coordination shell of the cluster, no conclusive answer can be given, how the different behavior of the proteins can be explained.

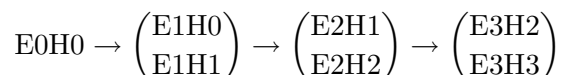


**Figure 10.6** Stereoview of an overlay of FeMoco and its environment in the resting state of *K. pneumoniae* (black [23]) and *A. vinelandii* (red [19]).

### 10.2.3 Determination of the chemical potential of protons

The protonation state of the resting state provides an upper bound for the proton chemical potential  $\mu_p$ . The fact that the first reduced state E1 is protonated in one species' enzyme and non-protonated in another's indicates that  $\mu_p$  is approximately the first protonation energy of E1. It can be expected that  $\mu_p$  only differs slightly between the two proteins as they exhibit similar reaction rates.

Thus, following Fig. 10.4 on page 124 the protonation state can be determined for each charge state to within one proton:



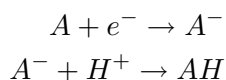
The notation  $\text{ExHy}$  describes a cofactor with  $x$  electrons and  $y$  protons added to the resting state. Each arrow denotes one electron transfer.

## 10.3 The chemical potential of hydrogen atoms

From the last section it follows that each reduction step induces one protonation step in the cofactor without  $\text{N}_2$  bound. Here I describe the procedure how the different numbers of electrons and protons enter the energetics of nitrogen conversion.

### 10.3.1 Evaluation of protonation energies

Nitrogen is converted to ammonia by adding electrons and protons to its complex with the cofactor. If I denote the cofactor before some protonation step with  $A$ , I need to consider two reaction steps: the reduction and the protonation:



The electrons and protons are transferred from a reservoir, namely the protein surrounding, to FeMoco. In order to discuss the energetics it is necessary to understand the chemical potentials of the reservoir for protons and electrons. The chemical potential for a particle type is the energy required to remove one particle from the reservoir. It is also the driving force to add this particle to the FeMoco. Only, if the chemical potential of the reservoir is larger than the energy required to add the particle to the cofactor, the particle will be transferred.

The electrons are supplied by the Fe-protein under MgATP hydrolysis as explained in section 3.1 on page 17. Association of the Fe-protein and the MoFe-protein induces electron transfer from the Fe-protein onto the P-cluster of the

MoFe-protein. From there, the electron is transferred to FeMoco. The energy of the electron just before the electron transfer is the ionization potential of the charged P-cluster. The relevant chemical potential  $\mu_e$  for the electron is therefore the ionization potential of the charged P-cluster embedded in the protein. Immediately following the transfer to the FeMoco, some energy is dissipated as the FeMoco adjusts its structure to the new charge state. The energy dissipated following the electron transfer is denoted as  $W$ . Thus the energy of the reduced cofactor is  $E[A^-] = E[A] + \mu_e - W$ .

The protons are transferred from the proton-transfer channel to the cofactor. Using the proton chemical potential of the channel ( $\mu_p$ ), the energy of the protonated cofactor is then  $E[AH] = E[A^-] + \mu_p = E[A] + \mu_e - W + \mu_p$ .

For the reduction and protonation steps leading from the resting state to the docking state of  $N_2$ , I found that electrons and protons are added one-by-one in a ping-pong like manner, see section 10.2.3. In the following, I assume the same behavior also for the remaining steps of the catalytic conversion.

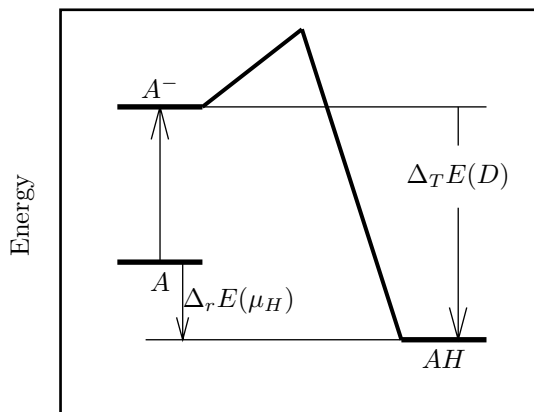
As it is not possible to directly derive  $\mu_e$  from the calculations, I define a single chemical potential for hydrogen atoms  $\mu_H = \mu_e - W + \mu_p$  which can be estimated.

Note, that the relaxation energy  $W$  may vary from step to step in the catalytic cycle. Nevertheless, I treat  $\mu_H$  as constant, which is justified by the small magnitude of the relaxation energy verified for one of the systems. Thus the reaction energy of a reaction  $A + e^- + H^+ \rightarrow AH$  is  $\Delta_r E = E[AH] - E[A] - \mu_H$ .

Currently, a quantitative determination of  $\mu_H$  from theory alone is not possible. Nevertheless I can obtain an estimate from experiment using the FeMoco as test probe. It is known from experiment that on the one hand protons are transferred to the resting state  $M$  under turnover conditions. Therefore,  $\mu_H$  is sufficiently high to drive protonation, that is  $\mu_H > E[MH] - E[M]$ . On the other hand, no protonation occurs under the same conditions, but in the absence of MgATP. Thus the chemical potential in the absence of MgATP, denoted by  $\mu'_H$ , must be sufficiently low not to drive protonation, that is  $\mu'_H < E[MH] - E[M]$ . As two MgATP are hydrolyzed in each electron transfer, the difference between the chemical potentials with and without MgATP is smaller than twice the hydrolyzation energy of MgATP [91], that is  $\mu_H - \mu'_H < 64.4 \text{ kJ/mol}$ . It is smaller, because a fraction of the energy supplied by MgATP will be dissipated. Therefore, I assume the lower bound for  $\mu_H$ , that is  $\mu_H = E[MH] - E[M]$  in my calculations. This is the most conservative assumption possible.

Previous studies [62–64, 104, 163] used  $\mu_H = \frac{1}{2}E(H_2)$  as proton chemical potential. It is 35.5 kJ/mol lower than my choice.  $\mu_H = \frac{1}{2}E(H_2)$  would be the appropriate choice if the hydrogen atoms would be obtained from gaseous hydrogen. While in this choice, the production of gaseous hydrogen  $2H^+ + 2e^- \rightarrow H_2$  is energetically neutral, this reaction is exothermic by 71 kJ/mol ( $= E(H_2) - 2\mu_H$ ) when using my choice of  $\mu_H$ .

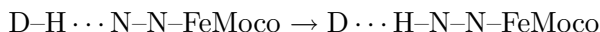
### 10.3.2 $\text{NH}_4^+$ as model for the proton source



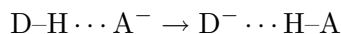
**Figure 10.7** Energy contributions for one reduction-and-protonation step  $A + e^- + H^+ \rightarrow AH$ . The reaction energy  $\Delta_r E$  only depends on the reaction at the cofactor and the hydrogen chemical potential  $\mu_H$  but is independent of the proton donor.  $\Delta_T E$  is the energy change of transferring a proton from the donor to the cofactor. As this step preserves the number of particles in the system,  $\Delta_T E$  can directly be calculated as long as both protonated and de-protonated states are minima.  $\Delta_T E$  is independent of  $\mu_H$ .

The energy of reduction and protonation, evaluated as described in the previous section as  $\Delta_r E = E[AH] - E[A] - \mu_H$ , only depends on the energies of the relaxed structures  $A$  and  $AH$  and the chemical potential  $\mu_H$ . However, protonation may induce structural changes between  $A^-$  and  $AH$ , such as forming or breaking bonds. These and the proton transfer itself result in a barrier for protonation. In order to estimate this barrier, I modeled the protonation reaction.

Protons are obtained from water in the bacterial cell and transported via a proton channel to the cofactor. The immediate protonation occurs within a hydrogen bond from a donor (D) to  $\text{N}_2$ , as in



or in the short-hand notation:



The crucial quantity, deciding if this proton transfer is possible, is the acidity of this donor. In general, the donor  $D$  will be a chain or a pool of water molecules containing one  $\text{H}_3\text{O}^+$ . In my calculations, I need to model the donor by isolated molecules. An isolated  $\text{H}_3\text{O}^+$  molecule is a poor model for the proton channel, because hydrolyzation significantly decreases the acidity of  $\text{H}_3\text{O}^+$ . Therefore I

have chosen an isolated  $\text{NH}_4^+$  ion, which has an acidity in between that of  $\text{H}_3\text{O}^+$  and  $\text{H}_2\text{O}$ .

As starting configuration for the proton transfer, I placed the  $\text{NH}_4^+$  ion near the expected position of protonation with a distance of 2 Å between the proton of  $\text{NH}_4^+$  and the atom, which is to be protonated. This distance corresponds to a typical hydrogen bond. For the proton transfer, all atoms are free to move. If the proton transfer does not occur spontaneously, the reaction has been investigated using a bond-length constraint between the donor and the proton.

Figure 10.7 describes the different energy contributions to the protonation energies. State  $A$  refers to the cofactor with bound substrate and  $AH$  refers to it after reduction and protonation. Their energy difference  $\Delta_r E = E[AH] - E[A] - \mu_H$  is independent of the proton donor.

The proton transfer energy  $\Delta_T E = E[AH \cdots \text{NH}_3] - E[A^- \cdots \text{NH}_4^+]$  requires two additional calculations. It is independent of  $\mu_H$  as the number of particles is preserved, but depends on the choice of the donor. Thus the energy difference between  $A$  and  $A^-$  does not only contain the reduction energy but also depends on both,  $\mu_H$  and the choice of the donor.

Note that the choice of the donor does not affect the energetics of the relevant intermediates of the reaction. It only affects the reduced state immediately before protonation and the barrier for protonation.

By choosing compounds  $A$ ,  $AH$ ,  $[A^- + \text{NH}_4^+]$ , and  $[AH + \text{NH}_3]$  with the same charge, I limit the errors caused by the lack of electrostatic screening by the environment.

The question remains, if or where suitable proton donors, which may form hydrogen bonds to  $\text{N}_2$ , are available in the protein. Such donors can be water molecules or amino acids occurring in two protonation states at physiological conditions. The only amino acid fulfilling this criterion is histidine. While there is no histidine located near the proton channel described in section 7.1.3 on page 85, several water molecules are present in the X-ray structure around the sulfur site bridging Fe3 and Fe7. As it will be shown in section 11.1.4 on page 143 these iron sites are best-suited for the catalytic reaction to take place. I conclude that the relevant proton donor is indeed  $\text{H}_3\text{O}^+$ .

## 10.4 Consequences of the protonation studies on the reaction mechanism

- ☞ The resting state is not protonated.
- ☞ Each reduction step induces a single protonation.
- ☞ Two protonations of the cofactor from the same proton path result in  $\text{H}_2$  formation.



## 11 Nitrogen reduction

*Galileo was no idiot. Only an idiot could believe that science requires martyrdom – that may be necessary in religion, but in time a scientific result will establish itself.*

David Hilbert

The nitrogen reduction mechanism, the main aim of this thesis, will be described in this chapter. Nitrogen binding induces openings of the cluster. Two binding modes play an important role in the conversion mechanism. Protons are added to lone-pairs at the nitrogen atoms which play the role of Brønsted bases, proton acceptors.

The notation for the structures follows the one chosen in paper 2. The resting state is denoted as  $M$ , which is reduced and protonated leading to  $MH$  and  $MH_2$ . The complexes following adsorption of dinitrogen are given letters in alphabetic order according to the number of proton transfers and numerals for their energetic order. A numeral 0 denotes the ground state for the selected composition. In case of the dinitrogen binding modes,

In this thesis, a larger number of modes are discussed than those mentioned in paper 2, therefore the labeling in section 11.1.1 differs from energetic order.

A summary of the nitrogen conversion mechanism will be given in chapter 13 on page 179.

### 11.1 Nitrogen binding

The first question to answer and a major key for unraveling the reaction mechanism is finding the site at which dinitrogen binds to the cofactor. In contrast to the protonation modes of the cluster, it cannot be addressed by investigating channels for  $N_2$  through the protein to the cluster. On the one hand, dinitrogen is very inert, uncharged, and apolar and thus shows hardly any interaction with the protein. Therefore it is very difficult to investigate a path of  $N_2$  through the protein to FeMoco. On the other hand, these weak interactions with the environment would allow  $N_2$  to move around the cofactor and reach its optimal binding site. Therefore, all possible  $N_2$  binding sites have to be investigated and the lowest-energy structure represents the real binding site.

The binding energy of dinitrogen depends on the reduction and protonation state of the cluster. Binding to the cluster in its E1H1 state is endothermic, but becomes exothermic with increasing number of electrons and protons donated to the cofactor.

### 11.1.1 Nitrogen binding modes

Several binding modes have been proposed in the literature for N<sub>2</sub> binding to the vacant cluster. Some of them are still possible with the central ligand included into the structural model, others are destabilized by the central ligand. However, there are also binding modes which only exist because of the central ligand. I will show that the only relevant binding modes, axial binding to one iron site (A0) and bridged binding between two iron sites (A1) are only possible in presence of the central ligand.

I investigated the binding energy and stability of the following N<sub>2</sub> binding modes. The stable modes are depicted in Fig. 11.1 on page 137 and their relative energies are compared in table 11.1 on page 138. The spin orderings used below have been explained in Fig. 9.6 on page 108.

#### Axial (head-on) coordinations

1. **Axial coordination to one Fe site next to Mo with opening of the SH-bridge (A0):** This is the lowest-energy structure. Binding of N<sub>2</sub> to one Fe site next to Mo induces an opening of the neighboring SH-group. It leaves all iron sites in a tetrahedral coordination and high-spin mode. The total spin of the system is  $S = 5/2$  with spin ordering BS7. The barrier for N<sub>2</sub> binding is 27 kJ/mol and the binding energy is -30 kJ/mol, thus binding is exothermic.
2. **Axial coordination to one Fe site next to Fe<sub>1</sub> with opening of the SH-bridge:** The structure is similar to (1.), the mode with binding next to Mo. Binding is less stable, -16 kJ/mol. The total spin is  $S = 1/2$  with a spin ordering BS7.
3. **Axial coordination to one Fe next to Fe<sub>1</sub> without opening of an SH-bridge:** Axial connection of N<sub>2</sub> to one iron site is also possible without opening of the corresponding sulfur bridge. This pulls the iron site, to which N<sub>2</sub> is coordinated, out of the cage and breaks its bond to the central ligand. Its tetrahedral environment is preserved. This structure is the starting point of the mechanism discussed by Hinnemann [104]. However, its energy is higher than (2.), exhibiting a broken sulfur bridge. The binding energy is -13 kJ/mol. Spin state:  $S = 5/2$ , BS6. The binding energy varies if



the participating iron atom is not fully pulled out of the cage: if the atom stays trigonal-bipyramidally coordinated, the binding energy is +8 kJ/mol,  $S = 3/2$ , BS6.

4. **Axial coordination to one Fe next to Mo without opening of an SH-bridge:** This is equivalent to the structure (3.) discussed above, but  $N_2$  is bound to an iron site next to Mo. This site is not fully pulled out of the cage. The binding energy is +26 kJ/mol. The spin structure is non-collinear,  $S = 1.07$ .

#### Bridged coordinations

5.  **$N_2$  coordination bridging between two Fe sites, the former SH-bridge is connected to Fe next to  $Fe_1$  (A1):** In this mode, which is part of the main reaction mechanism,  $N_2$  replaces the former SH-bridge between two Fe sites. The mode can be reached from mode (1.) after overcoming a barrier of 66 kJ/mol. The former SH-bridge is now connected to an Fe site next to  $Fe_1$ . The bond of that site to the central ligand is broken. Thus the approximate tetrahedral environment of all iron atoms is preserved as it can be seen from Fig. 11.1. The binding energy is -24 kJ/mol; the spin state is:  $S = 5/2$ , BS6. The structure stayed stable at a molecular dynamics run of 1 ps at 300 K.
6.  **$N_2$  coordination bridging between two iron sites, the former SH-bridge connected to Fe next to Mo:** This structure is equivalent to the one discussed above, but the SH-group is connected to an Fe site next to Mo. It can be reached via mode (2.). The binding energy is 0 kJ/mol. The spin structure is collinear,  $S = 1/2$ , BS6.

#### Head-on equatorial coordinations

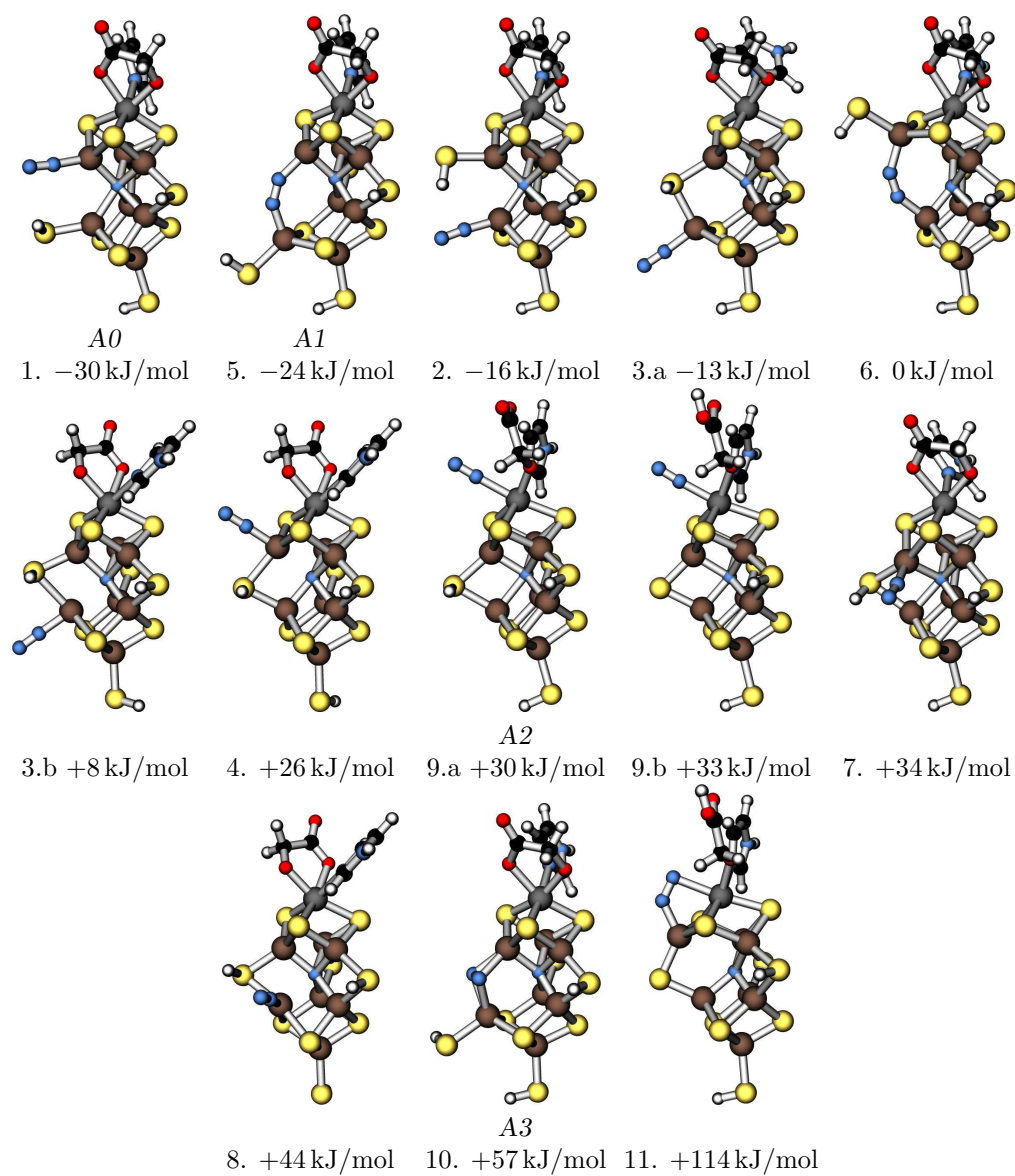
7. **Equatorial coordination to one Fe site next to Mo:** In this structure,  $N_2$  is bound equatorial to one Fe site. This site is coordinated approximately octahedral with one ligand missing. I found the corresponding mode to be the most stable  $N_2$  binding mode to the vacant cofactor, which is heavily distorted in that case. However, with central ligand, binding is endothermic due to the additional rigidity of the cage, +34 kJ/mol. All iron sites are high-spin, however, the spin of the Fe site connected to  $N_2$  is smaller than the spin of the other Fe sites.  $S = 3/2$ , BS6.
8. **Equatorial coordination to one Fe site next to  $Fe_1$ :** The structure is equivalent to (7.), the binding energy is +44 kJ/mol. The spin structure is nearly collinear,  $S = 1.24$ , similar to BS6.

**Other coordinations**

9. **Coordination to Mo (A2):** Coordination of  $N_2$  to the molybdenum site has been widely discussed in the literature [66–70, 99, 108, 164]. It requires a monodentate homocitrate ligand, one oxygen atom of homocitrate has to dissociate from Mo. Binding is endothermic, +30 kJ/mol. The spin structure is  $S = 5/2$ , BS6. Protonation of the carboxyl group of homocitrate, not connected to Mo any more makes the binding even slightly more unstable: +33 kJ/mol. The proton is taken from one of the sulfur bridges in order to keep the participating number of atoms fixed.
10. **Both nitrogen atoms bridging between two Fe sites (A3):** In this structure, both nitrogen atoms are connected to both iron sites. The iron d-orbitals stabilize the  $\pi^*$ -orbital of the N–N bond. The structure is metastable with a binding energy of +57 kJ/mol,  $S = 5/2$ , BS6.
11.  **$N_2$  bridging between Fe and Mo:** In this structure all sulfur bridges remain, but  $N_2$  coordinates to both Mo and one of its neighboring iron atoms. A structure similar to that one has been proposed by Durrant [69] after one protonation of  $N_2$ . The structure is a metastable minimum but with rather high energy: +114 kJ/mol,  $S = 5/2$ , BS7.
12.  **$\eta_2$ -binding to one Fe site:** A structure with both nitrogen atoms connected side-on to the same Fe site has been proposed by Dance [24]. It is no minimum in my calculations, neither starting from the experimental X-ray structure nor from axial coordination to Fe and tilting of bound  $N_2$ .  $N_2$  drifts away.
13. **Coordination with only one N atom bridging two Fe sites:** Corresponding structures with one nitrogen atom bridging both iron sites exist in higher reduced parts of the reaction mechanism as it will be described in the following sections. The other nitrogen atom points out of the cage and is not connected to any iron site. While this structure is stable after one protonation of  $N_2$ , it transforms into the axial coordination mode (1.) without proton addition.

Other binding modes, such as  $N_2$  above one face spanned by four iron atoms and connected to all four of them, or the slightly asymmetric version with one nitrogen atom in the center of the face, are not stable after inclusion of the central ligand.

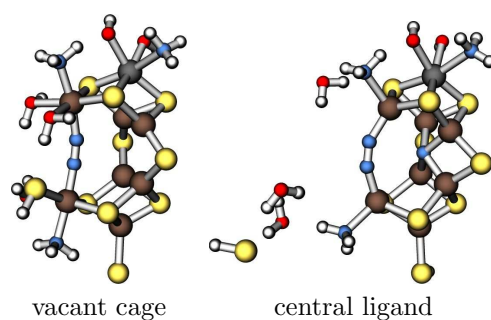
Nitrogen bridging between two iron sites has already been proposed by Sellmann et al. [165–168] in analogy to smaller Fe-complexes. In contrast to my model, they proposed the participating iron sites to be in an octahedral environment and



**Figure 11.1** Dinitrogen binding modes and the corresponding binding energy in energetic order.

state	$\Delta E_{\text{bind}}$ (kJ/mol)	$S$	spin ordering
1. ( <i>A0</i> )	−30	5/2	BS7
5. ( <i>A1</i> )	−24	5/2	BS6
2.	−16	1/2	BS7
3.a tetrahedral	−13	5/2	BS6
6.	0	1/2	BS6
3.b trig.-bipy.	+8	3/2	BS6
4.	+26	1.07	
9.a COO <sup>−</sup> ( <i>A2</i> )	+30	5/2	BS6
9.b COOH	+33	0.45	
7.	+34	3/2	BS6
8.	+44	1.24	
10. ( <i>A3</i> )	+57	5/2	BS6
11.	+114	5/2	BS7

**Table 11.1** Dinitrogen binding modes to the cofactor after two reduction and protonation steps (E2H2) in energetic order. Binding energy, total spin, and spin ordering of the bound states are given. The labeling of the states refers to the text.



**Figure 11.2** The model proposed by Sellmann and its destabilization by the central ligand.

low-spin. The additional ligands for the iron atoms should be water molecules and nearby protein residues. The low-spin configuration would lead to a stronger nitrogen binding.

A model according to that came out to be stable in my calculations without central ligand. The resulting structure can be seen in Fig. 11.2. However, including the central nitrogen ligand in the above calculated structure between the four remaining Fe sites leads to dissociation of the water ligands and the SH group and to a destabilization of the whole structure. It can be seen in Fig. 11.2 that the iron atoms formerly being octahedrally coordinated change into tetrahedral coordination. They also change into the high-spin state.

### Dependence of the results on approximations

**Unit cell:** The above given binding energies of the 15 considered binding modes have been calculated with different unit cells. However, each binding energy is the difference in total energy of two calculations with the same unit cell. Therefore the state without bound  $\text{N}_2$  has been calculated in different cells according to the size of the adducts. The binding energies only depend slightly on the size of the cell with differences of around 7 kJ/mol when different cells are used.

**Semicore states:** Inclusion of semicore electrons (Fe: 3s, 3p; Mo: 4s, 4p) of the metal atoms into the valence formalism weakens the bindings a little without changing the relative order of the binding modes. I calculated the most important modes, i.e. 1., 5., 4., and 9.a including semicore states for the metal atoms. The resulting binding energies have been used in the publication of these results in paper 1 [27] and are compared to the values obtained with usual accuracy in table 11.2.

state	$\Delta E_{\text{val}}$	$\Delta E_{\text{semicore}}$
1.	-30	-19
5.	-24	-14
4.	+26	+12
9.a $\text{COO}^-$	+30	+30

**Table 11.2** Binding energy of selected  $\text{N}_2$  binding modes when metal semicore states are treated as valence states. All energies are given in kJ/mol.

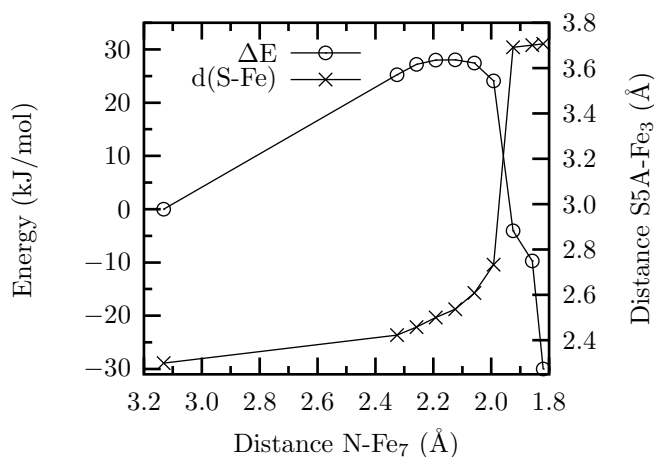
**XC-Functional:** The density functional may also have an effect on the binding energies. Mortensen et al. [169] obtained very accurate values for the binding of  $\text{N}_2$  to metal surfaces when using the revised PBE (RPBE) functional [48]. It generally lowers binding energies and enlarges bonds. Thus it always increases

the accuracy compared to PBE, which I used, whenever PBE overestimates bond strength. To enable comparison, I implemented the RPBE functional into the CP-PAW code and recalculated the most important binding energies with this functional. The results are given in table 11.3. As it can be expected, the binding is weaker, but the relative order of the modes is preserved and thus the same physics results from it.

state	$\Delta E_{\text{PBE}}$	$\Delta E_{\text{RPBE}}$
1.	-30	-7
5.	-24	+3
9.a COO <sup>-</sup>	+30	+43

**Table 11.3** Comparison of the functionals PBE and RPBE for N<sub>2</sub> binding energies (kJ/mol).

### 11.1.2 Axial nitrogen binding (A0)



**Figure 11.3** Energy barrier of opening of the sulfur bridge (S5A) and axial coordination of N<sub>2</sub> to Fe<sub>7</sub> of FeMoco. The two reactions occur in a concerted mechanism.

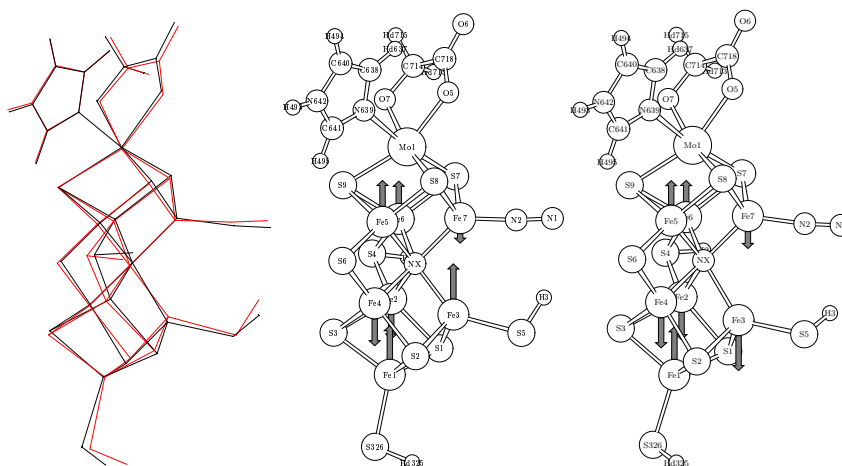
The first step in the mechanism, the interaction of the cofactor with N<sub>2</sub>, is axial head-on coordination of N<sub>2</sub> to one of the iron sites. This coordination induces an opening of the cluster and leads to binding mode A0 (1.). The barrier for binding N<sub>2</sub> from the gas-phase in that mode is 27 kJ/mol. Energetics of the adsorption process are illustrated in Fig. 11.3. N<sub>2</sub> coordination and opening of the sulfur bridge occur in a concerted mechanism.

It is unlikely that the closed axial structure (3. or 4.) is ever reached, when N<sub>2</sub> binds to the cluster. Opening of the sulfur bridge has a low barrier of < 10 kJ/mol

meaning that the closed structure is not stable.

In the open axial binding mode, two rather stable spin states have been found. The spin orderings and the corresponding atomic structures are compared in Fig. 11.4. The lower-energy structure exhibits  $S = 5/2$  (BS7) while the other state shows  $S = 3/2$  (BS6) and is 10 kJ/mol higher in energy with and without inclusion of semicore states. Both structures differ by a spin-flip of  $\text{Fe}_3$ . The  $S = 5/2$  ground state has been used in all comparisons to other structures.

Binding to the iron site weakens the N–N triple bond. The bond distance is increased from 1.105 Å in the vacuum to 1.141 Å in *A0*.



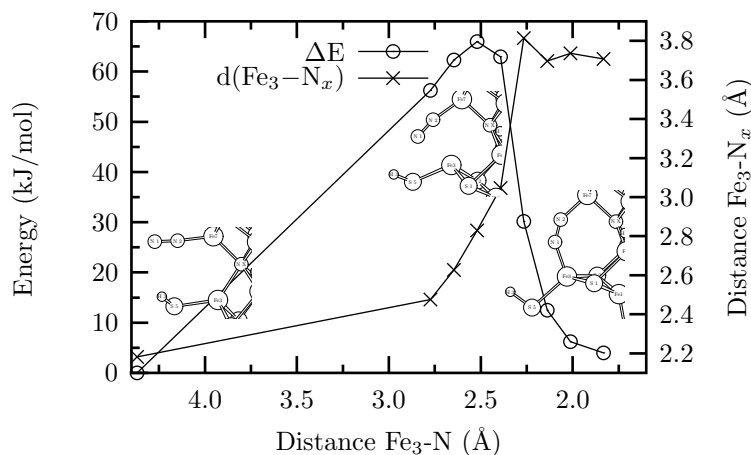
**Figure 11.4** The two spin states of the open axial structure. Left: a comparison of both structures (red: 5/2, black: 3/2). Middle: spin structure of the  $S = 5/2$  (BS7) state, Right: spin structure of the higher-energy  $S = 3/2$  (BS6) state. The major change is a spin-flip of  $\text{Fe}_3$ .

The open axial structure seems to be well suited for a direct protonation of  $\text{N}_2$  from the neighboring SH group as the H–N distance is only 3.0 Å. However, this proton transfer is not stable, the proton spontaneously moves back to the sulfur site.

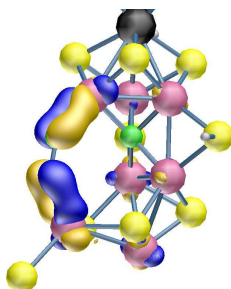
### 11.1.3 Bridged nitrogen binding (A1)

The bridged nitrogen binding mode *A1* (5.) is slightly less stable than the axial mode *A0*. The latter is the first coordination mode. However, after overcoming a barrier of 66 kJ/mol the system may change from *A0* to *A1*. Details of the energy surface of this reaction are given in Fig. 11.5. While the total spin is conserved in this transition, the spin ordering changes from BS7 to BS6, a spin-flip at  $\text{Fe}_3$ . Therefore, some intermediate state of the transition show non-collinear spin orderings and the barrier may only be calculated accurately with non-collinear

spin description. The barrier of 66 kJ/mol can be overcome at room temperature within the time scale provided by the reaction rate, see section 5.4.



**Figure 11.5** Transition from the axial to the bridged binding mode. The reaction coordinate is mapped onto the distance from  $\text{Fe}_3$  to the terminal nitrogen atom of  $\text{N}_2$ . The bond from  $\text{Fe}_3$  to the central ligand  $\text{N}_x$  breaks during this rearrangement. Its bond length is shown on the right ordinate. Start-structure, transition state and end structure are also shown.



**Figure 11.6**  $\pi$ -backdonation: the picture shows isosurfaces of an occupied orbital which weakens the N–N triple bond in the  $A1$  binding mode. It is a state of the minority-spin direction of both participating iron atoms.

One could have expected that the central ligand adds rigidity to the cofactor. However, the opposite is the case. The ligand is able to offer a varying number of bonds to its iron neighbors and thus enables binding modes like the bridged one.

In this binding mode, dinitrogen is already activated through the partial occupation of the  $\pi^*$ -orbital. This orbital, shown in Fig. 11.6, is populated by  $\pi$  bonds between dinitrogen and the iron atoms. The activation is also apparent in the enlargement of the N–N bond from 1.105 Å to 1.173 Å.



### 11.1.4 Embedding of bound nitrogen in the protein

#### Calculational details

A larger environment of the cofactor including protein residues is required to decide if there is enough space to accommodate  $N_2$  bound to the cofactor. Such a large model cannot be described by DFT calculations. Therefore I modeled the protein with the classical force field UFF (universal force field) [170]. It includes parameters of all atoms in the periodic table and allows efficient simulations of large systems. The energies obtained by this force field, however, are significantly less accurate than those obtained from DFT calculations.

To calculate embedding energies, I used the rigid model of the cofactor, obtained from DFT calculations, and inserted it into the molecular mechanics environment. The embedding energy is the energy difference between the force field energy of a structure with adsorbed  $N_2$  and the one of the bare cofactor. Bonds connecting quantum-mechanically described and classically described atoms have been saturated with hydrogen in the quantum (DFT) model. Atomic charges have not been considered. Axial and bridged binding to all six of the prismatic iron sites have been investigated.

As bound  $N_2$  requires some space in the protein, the embedding energies of  $N_2$  adsorption modes are rather high. However, due to the inaccuracy of the force field, differences between the embedding energies of different modes are more significant than absolute energies.

#### Embedding energies

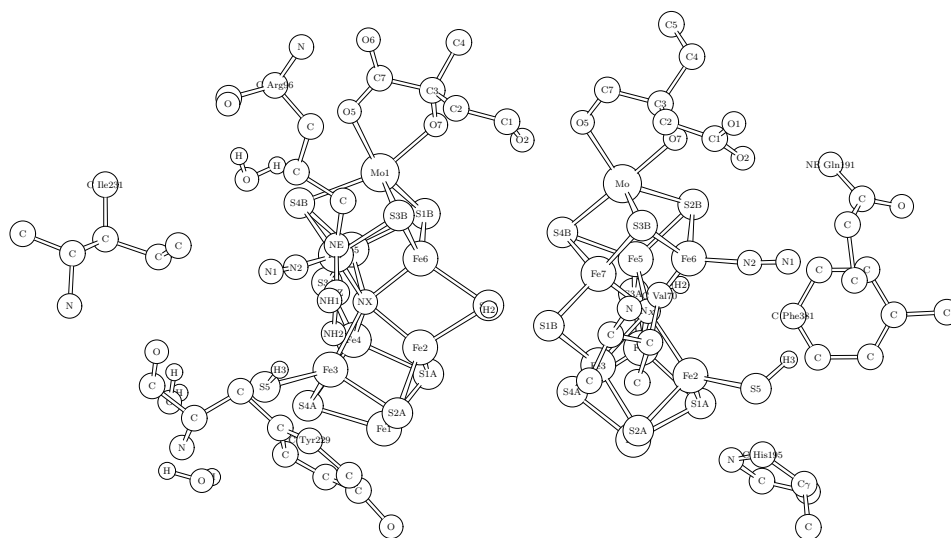
mode	N	S	kJ/mol
axial	7	3	160
bridged	7	3	60
axial	3	7	160
bridged	3	7	100
axial	6	2	260
bridged	6	2	260
axial	2	6	300
bridged	2	6	220
axial	5	4	220
bridged	5	4	430
axial	4	5	350
bridged	4	5	230

**Table 11.4** Embedding energies of different  $N_2$  binding modes. The second column gives the number of the Fe site to which  $N_2$  is bound, the third column gives the Fe site, the SH group is bound to.

The resulting embedding energies are given in table 11.4. It can be seen that

embedding can most easily be accommodated at the iron sites 3 and 7 in both the axial and the bridged binding mode. A transition between the binding modes is also possible because there is enough space around the SH group bound to Fe<sub>3</sub> or Fe<sub>7</sub>. Both, N<sub>2</sub>-binding to Fe<sub>3</sub> and to Fe<sub>7</sub>, are possible.

N<sub>2</sub> coordination to the iron sites 6 and 2 has a significantly higher embedding energy, thus I consider it rather unlikely that coordination to one of them takes place. If nitrogen initially binds to Fe<sub>6</sub> in the axial mode, a transition to the bridged mode is not possible because of steric hindrance by the His $\alpha$ 195 residue as it can be seen in Fig. 11.7.



**Figure 11.7** The environment of the axial mode with N<sub>2</sub> bound to Fe<sub>7</sub> and to Fe<sub>6</sub> after relaxation of the environment. Some H sites are omitted in the graph. Note, that the labels of the sulfur atoms are not properly adapted to the 1M1N nomenclature.

Although the embedding energy for N<sub>2</sub> binding to Fe<sub>4</sub> or Fe<sub>5</sub> is not too high in some modes, I consider it as unlikely because in these cases, a protein backbone chain has to move, which is expected to require more energy in reality than in the force-field simulation.

Thus, concerning the embedding, nitrogen binds to Fe<sub>3</sub> or Fe<sub>7</sub>. These sites are located in the direct vicinity of the only proton path capable of transporting more than one proton to the cofactor.

### 11.1.5 Dependence of the binding energy on the oxidation and protonation state of the cofactor

N<sub>2</sub> binding energy depends on the reduction and protonation state of the cluster as it can be seen from table 11.5. Generally the binding becomes stronger

with reduction of the cluster. This can be rationalized by the fact that the  $\pi^*$  orbital of dinitrogen has to be partially occupied in the bound state, which is easier if more electrons are available to the cluster. The differences between axial and bridged binding mode are small, actually they are below the accuracy of the methodology used.

While reduction strongly stabilizes binding, protonation weakly de-stabilizes bound  $N_2$ . It binds at all states considered in table 11.5 except E1H1, where the binding energy is too high for stable  $N_2$  coordination.

axial $N_2$	E1	E2	E3	bridged $N_2$	E1	E2	E3
H1	-9	-37		H1	-4	-34	
H2		-30	-40	H2		-24	-42

**Table 11.5** Binding energy of  $N_2$  (kJ/mol) at different oxidation- and protonation states of the cluster.

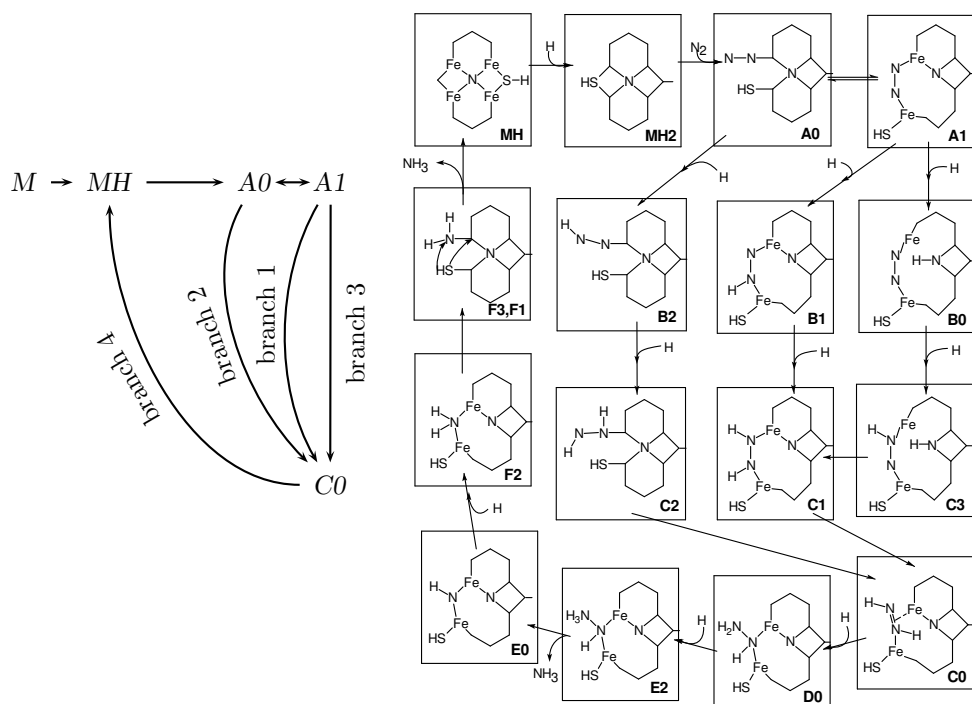
This is the reason why the cofactor of nitrogenase is not able to bind  $N_2$  at the resting state E0H0, it is not even able to bind it at E1H1, the first reduced state. The experimental fact, expressed in the Thorneley-Lowe scheme (see section 3.2 on page 18) that nitrogen binds after 3-4 reduction steps of the protein can be explained by the assumption that one electron transferred to the MoFe protein does not reach the cofactor but stays at the P-cluster. This assumption is supported by the experimental EPR observation [82], that one of the first electrons which are transferred to the MoFe-protein does not reach the FeMo-cofactor. Therefore the state E3H3 of the Thorneley-Lowe scheme corresponds to the state E2H2 in my analysis.

## 11.2 The reaction branches

From the two relevant dinitrogen binding modes, the axial and the bridged binding mode, the reaction cycle of nitrogen fixation at the cofactor of nitrogenase splits into different reaction branches which will be discussed in the following. They are outlined in Fig. 11.8 on the next page.

Starting from the resting state  $M$  ( $=E0H0$ ), dinitrogen binds after the first two reductions and protonations ( $MH2=E2H2$ ), as it has been discussed in section 11.1. From here, that is, from the axial binding mode  $A0$ , the cycle splits into three potential pathways, which all join after two reductions and protonations at  $C0$ , from which only one possible reaction path emerges. This branch 4 leads back to  $MH$ . Only one further reduction and protonation step is required before the next dinitrogen binds to form  $A0$ .

Therefore I divide the discussion of the reaction cycle into four different parts.



**Figure 11.8** Reaction scheme of the nitrogen conversion cycle.

In the first three parts I discuss the three branches of the catalytic cycle leading from  $A0$  to  $C0$ , namely

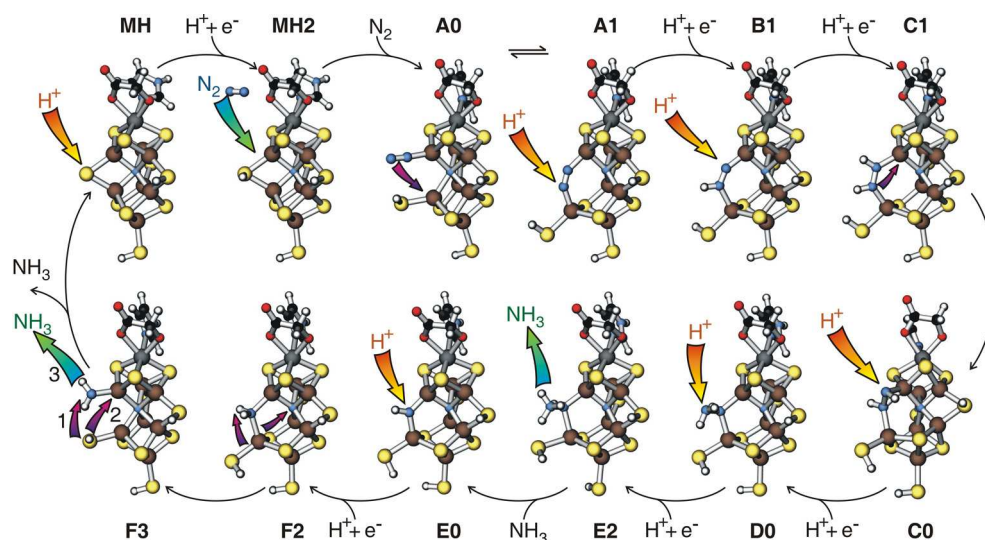
- Branch 1:  $A0 \rightarrow A1 \rightarrow B1 \rightarrow C1 \rightarrow C0$
- Branch 2:  $A0 \rightarrow B2 \rightarrow C2 \rightarrow C0$
- Branch 3:  $A0 \rightarrow A1 \rightarrow B0 \rightarrow C3 \rightarrow C1 \rightarrow C0$

Finally I discuss

- Branch 4:  $C0 \rightarrow D0 \rightarrow E2 \rightarrow E0 \rightarrow F2 \rightarrow F3 \rightarrow F1 \rightarrow F0/MH$ ,

which leads, via  $MH$  and  $MH2$ , back to  $A0$ . It will become clear in the following sections that the main reaction proceeds via the branches 1 and 4.

I calculated all intermediates and barriers of different branches of the nitrogen conversion cycle. The energies of the main branches, 1 and 4, are schematically shown in Fig. 11.10. The reaction rates  $\Gamma$  can be estimated as described in section 5.4. All barriers in the proposed cycle are sufficiently low so that all steps can proceed within the time scale of protein dissociation and association, that limit the turnover time.



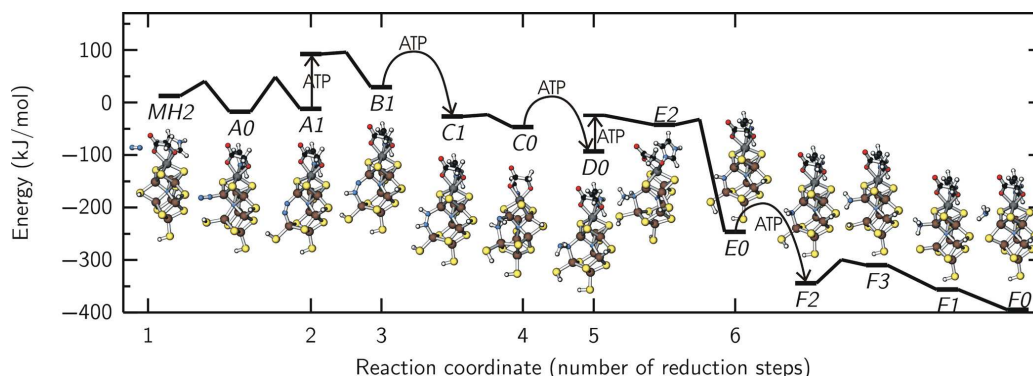
**Figure 11.9** Overview of structures and reactions of intermediates of the main path, branches 1 and 4, of the reaction mechanism.

### 11.2.1 Branch 1: Protonation of bridged dinitrogen (A1)

Here, the most favorable path from *A0* to *C0* will be discussed. As shown in section 11.1, the initial axial binding mode *A0* of dinitrogen and the bridged binding mode *A1* are nearly degenerate. My calculations yield the bridged mode marginally higher, that is by 6 kJ/mol, in energy. The transition from the axial to the bridged binding mode has a barrier of 66 kJ/mol. The rate for overcoming this barrier exceeds that of the electron transfer, which indicates that both binding modes *A0* and *A1* are in equilibrium. The reason for giving a slight preference to a cycle proceeding via the bridged binding mode will be provided below.

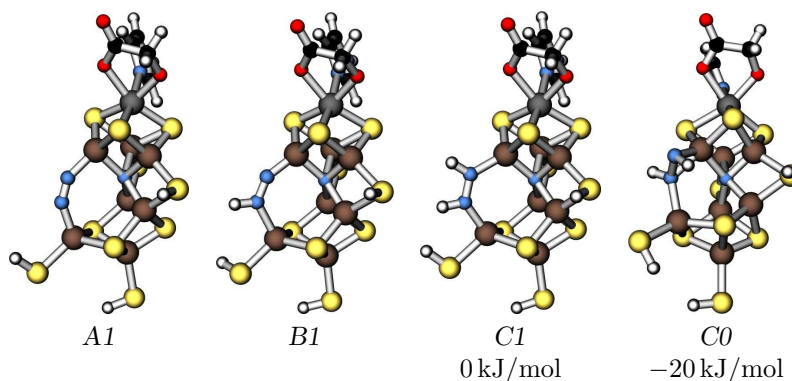
Fig. 11.10 shows a scheme of the energetics of the branches 1 and 4, the main path of the reaction. The vertical or curved arrows in Fig. 11.10 indicate reduction by one electron and subsequent protonation. After reduction, the proton is first added to an ammonia. If, on the one hand, the system with ammonium is metastable, a vertical arrow is drawn to the energy of the intermediate with the proton still residing on the donor. If, on the other hand, the proton transfer from ammonium to the substrate is spontaneous, a curved arrow is drawn to the resulting intermediate.

The energy required for the reduction and protonation steps, represented by the length of the vertical arrows, is supplied by the hydrolysis of MgATP. Its value is reflected in the chosen value of the chemical potential  $\mu_H$ . As discussed before, the value of the chemical potential is not known a-priori. I have chosen here the most



**Figure 11.10** Energy scheme of the total nitrogen conversion cycle. The arrows indicate protonations. Their lengths depends on the choice of  $\mu_H$ . Intermediates connected by bold lines contain the same number of atoms. The energy difference between these is independent of  $\mu_H$ .

conservative value compatible with experiment. My choice has been motivated in section 10.3, and the implication of different values for the chemical potential will be discussed in section 11.3.3.



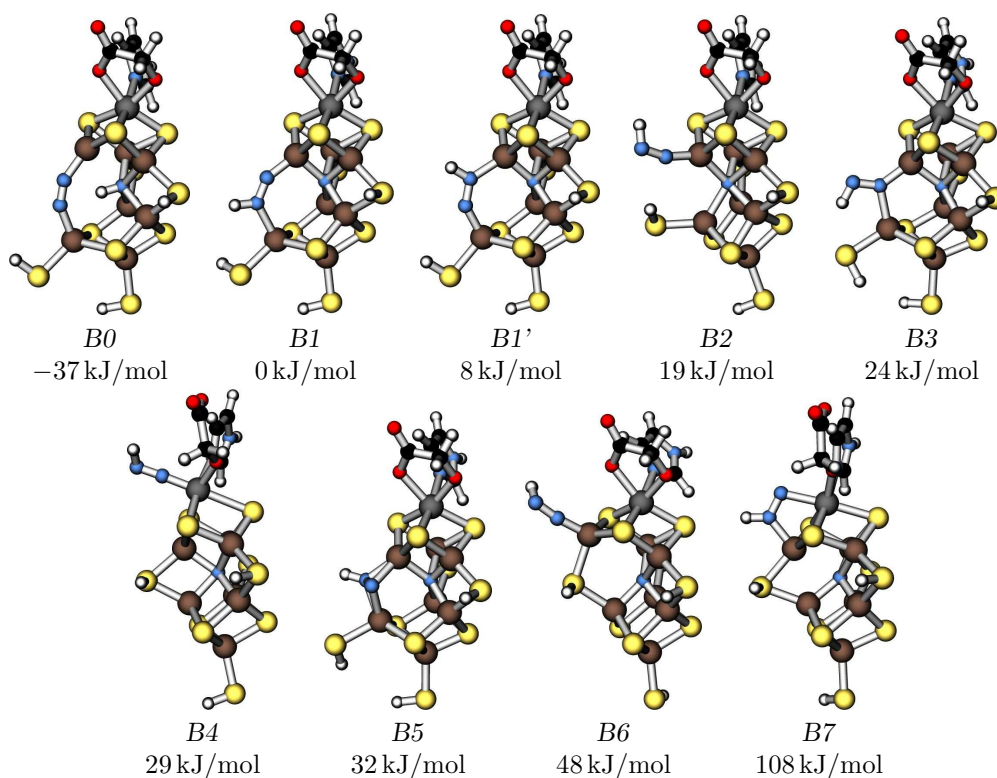
**Figure 11.11** Intermediates of branch 1: Protonation of bridged dinitrogen. The energy is given relative to  $C1$ .

**First protonation:** In the bridged mode, dinitrogen offers two proton acceptor sites, namely one on each nitrogen atom. The protonation energies of these two acceptor sites are comparable, their difference is only 8 kJ/mol. Protonation leads to structure  $B1$  or  $B1'$  shown in Fig. 11.11 and in Fig. 11.12. Reduction and protonation, taken together, are endothermic by 41 kJ/mol, using my choice of  $\mu_H$ .

In the vacuum reaction from  $\text{N}_2$  to  $\text{NH}_3$ , the first protonation is the most difficult step. Using the same choice for  $\mu_H$ , it requires an energy of 164 kJ/mol. Also in nitrogenase the first protonation is the most endothermic step, even though it requires much less energy than in vacuum, namely only 41 kJ/mol.

The cofactor activates dinitrogen by already forming two bonds to dinitrogen. These bonds result in a back-donation of electrons to the  $\pi^*$  orbitals, weakening the dinitrogen bond. Correspondingly, the dinitrogen bond is expanded from 1.105 Å in the vacuum to 1.173 Å in the bridged configuration *A1*.

Interestingly, the protonation from an ammonium, which mimics the proton donor, proceeds with a negligible barrier of only 4 kJ/mol and is exothermic by 63 kJ/mol. Note, however, that a protonation energy taken individually is not as reliable as the reaction energy, which is the sum of the reduction and protonation energies, as the former depends on the choice of ammonium as proton donor.



**Figure 11.12** Possible structures after the first protonation of  $\text{N}_2$  bound to FeMoco. The lowest-energy structure *B0* is not accessible because of a large barrier of 79 kJ/mol for protonating the central ligand as discussed in section 11.2.3. Therefore the energies are given relative to *B1*.

Because this protonation state of dinitrogen plays a special role as the ener-

getically most unfavorable intermediate of the reaction, I have calculated nine different structures for the complexes with  $\text{N}_2\text{H}$ , which are shown in Fig. 11.12.

- The energetically lowest isomer is *B0* with the proton not added to dinitrogen but to the central ligand. As discussed in section 11.2.3 this can be excluded due to a high barrier of 79 kJ/mol.
- The most stable configuration is the axial binding mode *A0*. Its protonation leads into the second branch described in section 11.2.2. While an unambiguous selection is not possible because the relevant energies are within my error bars, we consider the branch via *A1* more favorable than the one via protonation of *A0*, because the bridged mode becomes slightly more stable upon reduction by 7 kJ/mol. After protonation the bridged mode *B1* is 19 kJ/mol more stable than the axial mode *B2*.
- Only 4 kJ/mol above the energy of *B2*, I found the complex *B3* bridging two Fe sites with a single nitrogen atom. This structural principle is found again later in the reaction cycle as *D0* and *E2*. However, like the axial binding mode, it lies substantially, that is by 24 kJ/mol, higher in energy than the bridged mode *B1*.
- The protonated dinitrogen bound to molybdenum, i.e. *B4*, lies 29 kJ/mol above *B1*. It will be discussed in more detail in section 11.2.5, where I discuss the potential role of molybdenum.
- Dinitrogen can also bridge the two Fe sites with its axis perpendicular to the Fe-Fe direction, so that both nitrogen atoms are connected to both Fe sites. A similar binding mode has recently been found for nitrogen bridging two zirconium centers [171]. The energy of this structure, *B5*, lies 32 kJ/mol above that of *B1*.
- In the relevant intermediate of the mechanism proposed by Hinnemann and Nørskov [104], namely *B6*, dinitrogen binds axially to one Fe site like my structure *B2*. In contrast to *B2*, however, the sulfur bridge is still intact and the bond between this Fe site and the central ligand is broken. This structure is 48 kJ/mol above *B1* and it is 29 kJ/mol less stable than my axial binding mode *B2*.

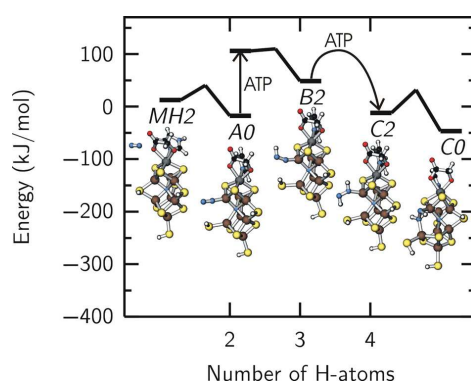
**Second protonation:** After one of the protonation sites has been saturated, there is only one protonation site left. Protonation of *B1* is spontaneous from  $\text{NH}_4^+$  and leads to structure *C1* shown in Fig. 11.11, containing the *cis*-form of the  $\text{H}-\text{N}=\text{N}-\text{H}$  adduct.



The *cis*-form is only metastable and converts into the more stable *trans*-form (*C0*). In this *trans*-form, dinitrogen forms a  $\pi$ -complex with one Fe site, while another Fe site binds saturates one of its lone pairs. The barrier of 26 kJ/mol for the conversion into the  $\pi$  complex can readily be overcome. The energy gain of this conversion is 20 kJ/mol.

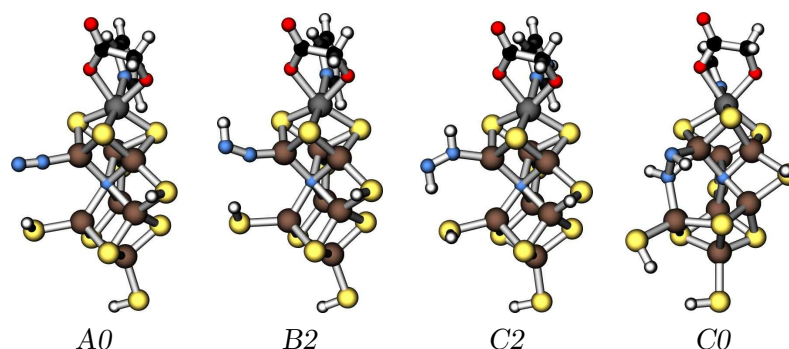
I will proceed with the reaction cycle after discussing the competing branches leading to *C0*.

### 11.2.2 Branch 2: Protonation of axial dinitrogen (*A0*)



**Figure 11.13** Energy scheme of branch 2: Protonation of axial dinitrogen (*A0*). See Fig. 11.10 for details.

The energetics of the first protonation steps starting from *A0* and leading to *C0* are depicted in Fig. 11.13.



**Figure 11.14** Intermediates of branch 2: Protonation of axial dinitrogen (*A0*).

**First protonation:** If  $\text{N}_2$  does not move into the bridging position but is axially bound (*A0*) during the first protonation, the proton will be added to the terminal nitrogen site. This results in structure *B2*. As seen in Fig. 11.14, the protonation does not induce major structural changes. The energy of the *B2* is 19 kJ/mol higher than *B1*, i.e. the bridged mode. I attribute the energy difference between *B1* and *B2* to the different occupation of the intra-plane  $\pi^*$  orbital of  $\text{N}_2$ , which is larger in *B1* due to the presence of the second Fe–N bond.

**Second protonation:** The second proton attaches to the proximal atom of the axially bound dinitrogen, i.e. *C2*, as shown in Fig. 11.14. This second protonation is spontaneous, using  $\text{NH}_4^+$  as donor. The second proton is added in a *trans*-position to the first. The Fe–N–N bond angle is close to  $120^\circ$ , as expected for  $\text{sp}^2$  hybridization.

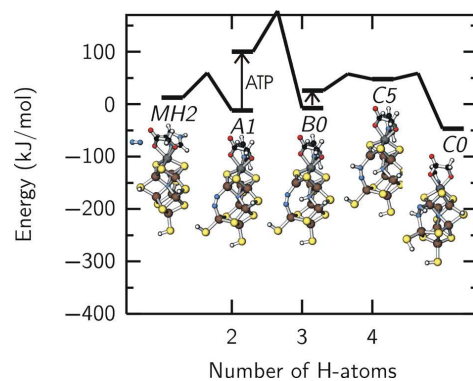
This state, *C2*, is only metastable, and transforms into structure *C0* gaining 35 kJ/mol. The barrier of this transformation is 44 kJ/mol.

**Side branch to *D0* bypassing *C0*:** Here I discuss a side branch, which may be employed, if the transformation from *C2* to *C0* is not completed before the next proton is added: if the next electron transfer proceeds, while the system still resides in *C2*, it can be directly protonated at the terminal nitrogen site resulting in *D2*. Intercalation of the proximal nitrogen atom between the two Fe sites leading to *D0* is exothermic by 19 kJ/mol. It requires only a small barrier of  $< 12$  kJ/mol. In *D0* the intercalated nitrogen atom is tetrahedrally coordinated to two Fe sites, one hydrogen and the  $\text{NH}_2$  group.

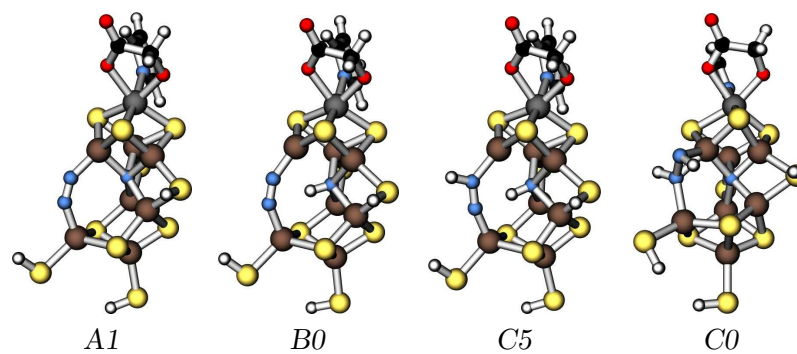
Here I rationalize why this intercalation from *D2* to *D0* proceeds with a lower barrier, i.e.  $< 12$  kJ/mol, than the formation of the  $\pi$ -complex *C0* from *C2*, which has a barrier of 44 kJ/mol: a precondition for the intercalation is the rotation of dinitrogen about the Fe–N bond. The interaction between the  $\pi$  orbitals of the double bonded  $\text{NH}=\text{NH}$  in *C2* with the Fe-d orbitals lock the rotation. The additional proton converts the double bond to a single bond, which frees the rotation and thus reduces the barrier.

### 11.2.3 Branch 3: Protonation of the central ligand

The third branch from *A0* to *C0* proceeds via protonation of the central ligand. I consider this branch unlikely, because it requires to overcome a large barrier of 79 kJ/mol. This barrier corresponds to a rate lower than the electron transfer rate. Nevertheless it is a viable path, if the electron transfer is slowed down. Furthermore, I must consider the errors caused by my theoretical description, that may affect the calculated barriers. The energetics of this branch is summarized in Fig. 11.15.



**Figure 11.15** Energy scheme of branch 3: Protonation of the central ligand. Protonation of the central ligand, A1 to B0 has a high barrier of 79 kJ/mol. The following first protonation of N<sub>2</sub>, leading to C5 is energetically unfavorable. See Fig. 11.10 for details.



**Figure 11.16** Intermediates of branch 3: Protonation of the central ligand.

**First protonation:** Nitrogen in the bridged binding mode *A1* opens the structure of the cofactor and leaves a lone-pair at the central ligand. Protonation of the central ligand results in *B0*, which is lower in energy by 37 kJ/mol than *B1*, the protonated dinitrogen bridge. There is only limited space for a donor to access the central ligand. The barrier for protonation from  $\text{NH}_4^+$  is 79 kJ/mol. This corresponds to a reaction rate of the order of  $0.1\text{--}1\text{ s}^{-1}$ , substantially lower than the electron transfer rate.

The relative energies of *B1* and *B0* may be affected by solvent effects not included in my calculations: I expect that the protein stabilizes *B1* more than *B0*. In *B1*, the proton at  $\text{N}_2$  may form hydrogen bonds with the environment, namely water or an arginine. I estimate this effect to be 10–30 kJ/mol, the typical hydrogen bond energy. The proton at the central ligand could not form such a hydrogen bond. No significant complex bonds between the three-coordinate Fe site and a water formed in my calculations: For *B0* an Fe–O distance of 2.7 Å has been calculated, which is substantially larger than typical complex bonds to Fe.

I attribute the barrier of 79 kJ/mol to a bond dissociation: upon protonation of the central ligand, the second of its bonds to iron sites breaks. As a result, one iron site is left in an unfavorable triangular coordination.

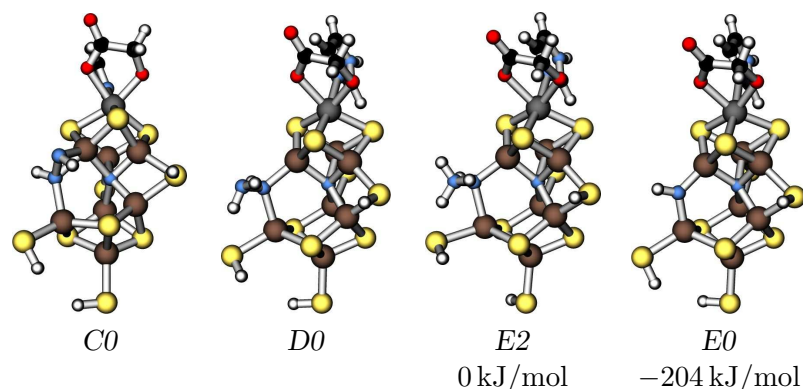
One may ask again, if the solvent effects affect the barrier. I find that the interaction between an additional  $\text{H}_2\text{O}$  molecule and the triangular Fe site in the initial state *A1* and in the transition state is even weaker than in the final state *B0*. Thus I conclude that the large barrier is not strongly affected by the solvent.

**Second protonation:** In state *B0*, dinitrogen is still unprotonated. The next proton is transferred to dinitrogen leading to *C5*, since I found that the central ligand does not accept further protons. As seen in Fig. 11.15, however, this proton transfer does not proceed easily: the intermediate *C5* is energetically fairly unfavorable. The proton transfer from ammonium is even endothermic. The reason for this difficulty is that the reduction step preceding protonation does not activate the N–N bond, as can be seen both from the  $\text{N}_2$  bond length and from a population analysis.

From *C5* the proton of the central ligand moves to dinitrogen leading to the intermediate *C0*. The barrier of 20 kJ/mol can readily be overcome and the reaction is exothermic by 69 kJ/mol.

#### 11.2.4 Branch 4: From *C0* to ammonia release

All three branches starting from axially bound dinitrogen join in the common structure *C0*. In *C0*, the  $\text{H–N=N–H}$  adduct forms a  $\pi$ -complex with one iron site and a  $\sigma$ -bond to another Fe site as seen in Fig. 11.17.



**Figure 11.17** First four intermediates of branch 4: from *C0* to ammonia release. Protonation of state *D0* leads to a high-energy intermediate state *E2*. Ammonia is released after overcoming a small barrier reaching *E0*. The energy is given relative to *E2*.

**Third protonation:** *C0* has only a single proton acceptor site on dinitrogen, namely *trans* to the Fe–N  $\sigma$ -bond.

Following reduction and protonation, *C0* converts into the  $\sigma$ -complex *D0*. Here, one nitrogen forms an  $\text{NH}_2$  group connected to the other one, which is tetrahedrally coordinated to two Fe-atoms, one H atom and the first N atom. During this protonation, the N–N double bond is converted into a single bond.

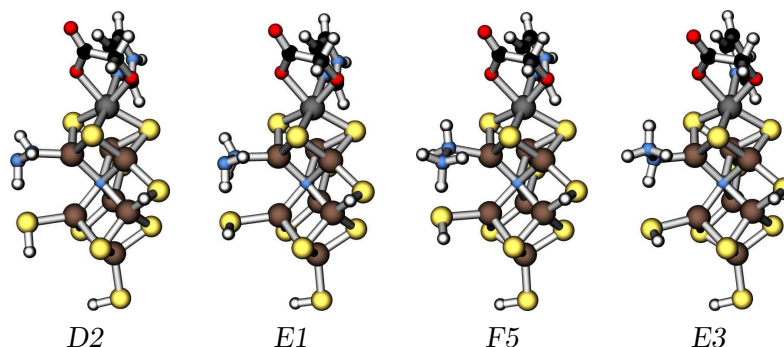
From branch 1, the intermediate *C0* can be bypassed: *C1* can directly be converted into *D0*. This reaction, however, requires that the system is reduced before the barrier from *C1* to *C0* has been overcome. Thus it will occur only at high electron flux. In this case, the third proton is added to either of the two nearly equivalent nitrogen atoms of *C1*. The reaction of the protonated *C1* to *D0* is exothermic by 18 kJ/mol.

**Fourth protonation:** Protonation of the only lone pair of structure *D0*, namely the one at the terminal  $\text{NH}_2$  group, requires only a small barrier of 11 kJ/mol using ammonium as proton donor. It results in a rather long but still intact N–N bond in structure *E2*.

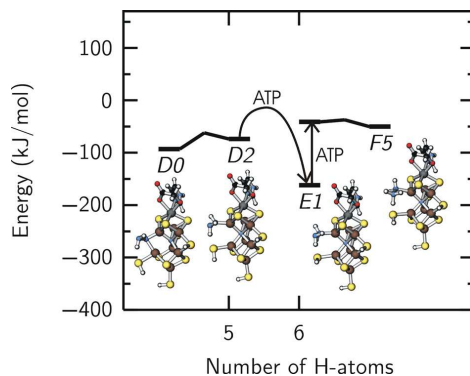
From intermediate *E2*, shown in Fig. 11.17, an ammonia molecule is easily dissociated, leading to structure *E0*. The barrier to cleave the N–N bond is only < 10 kJ/mol, and the reaction is strongly exothermic by 204 kJ/mol.

The mechanism of breaking the dinitrogen bond is interesting and provides an indication, why nature has chosen a complex cofactor such as the FeMoco. Therefore I discuss this step here in more detail. In order to break the N–N single bond, an additional electron pair has to be supplied by the cluster to dinitrogen. Thus, two iron sites have to be reduced simultaneously. (1) This is facilitated

by a large cluster with a delocalized electron system, because the change of the average oxidation state is smaller than in a small metal cluster. (2) The electron pair consists of one spin-up and one spin-down electron. A single high-spin Fe atom can only provide electrons with parallel spin. In the FeMoco, dinitrogen is in direct contact with two Fe sites with antiparallel spins, which can supply an electron pair with antiparallel spin.



**Figure 11.18** Intermediates of an alternative route leading to cleavage of the N–N bond.



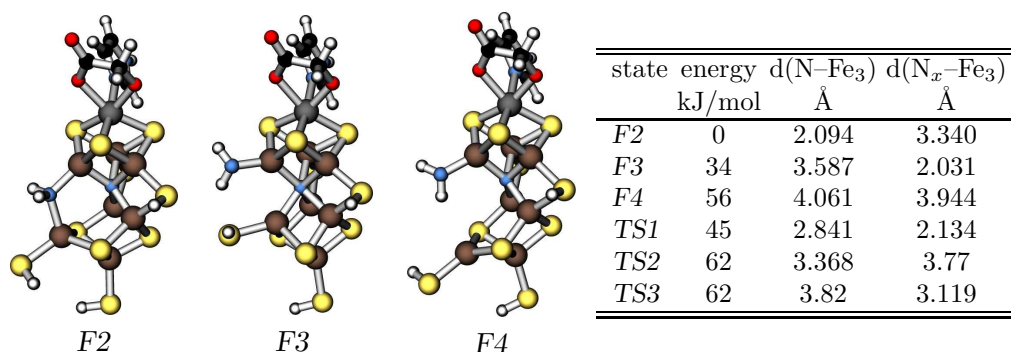
**Figure 11.19** Energy scheme of an alternative route leading to cleavage of the N–N bond. As protonation of *E1* is energetically unfavorable, the system would be trapped in *E1* and possibly react back. See Fig. 11.10 for details.

There is an alternative route shown in Fig. 11.19 leading to cleavage of the N–N bond. While I consider this side branch as unfavorable, it shall be described here for sake of completeness, because this alternative side branch is analogous to a pathway suggested earlier [62–64, 104]. Differing from that work, however, I consider here the SH bridge to be open. Starting from *D0*, the system can break one Fe–N bond and convert into *D2*, having an axially bound NH–NH<sub>2</sub> as shown

in Fig. 11.18. *D2* is 19 kJ/mol higher in energy than *D0* and therefore populated with lower probability.

Reduction and protonation of the *D2* at the proximal nitrogen atom leads to *E1* shown in Fig. 11.18, a hydrazine bound head-on to one Fe site. Protonation at the terminal hydrogen leading to *E3* can be excluded because it requires 138 kJ/mol more energy than protonation at the proximal nitrogen leading to *E1*. In *E1*, however the system is trapped: the next reduction and protonation step leads to the unfavorable intermediate *F5* with a hydrazine axially bound to the Fe site. The additional proton, however, is not stable on hydrazine in the presence of an ammonia. It can only be stabilized, if ammonia is prevented from forming any other hydrogen bonds. Therefore I expect the system to be trapped for a long time in *E1* and to react back via *D2* to *D0*.

With the exception of *E1* all intermediates of this side branch are higher in energy than those of the main branch. *E1* is 120 kJ/mol more stable than *E2* and 84 kJ/mol less stable than *E0* into which *E2* transforms. The intermediate *F5* of the side branch, however, is 294 kJ/mol higher in energy than *F2* of the main branch. Thus I conclude that this branch is not relevant.



**Figure 11.20** Energies relative to *F2* and geometries of intermediates and transition states involved in a rearrangement after the fifth protonation.

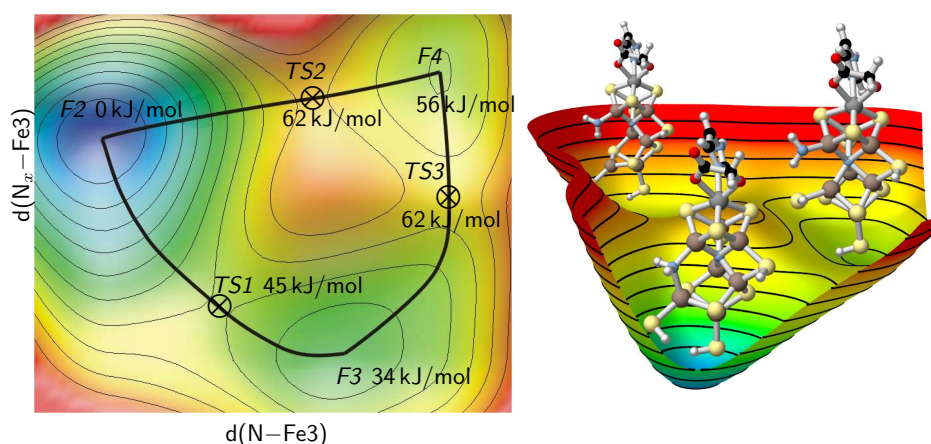
**Fifth protonation:** Protonation of the intermediate *E0* at the bridging NH group is spontaneous, using ammonium as proton donor, and leads to *F2*.

*F2* does not have proton acceptor sites at nitrogen. In order to expose a lone pair, one Fe-N must break, leading to *F3*. While this Fe-N bond is broken, one Fe-N bond to the central ligand is restored, so that the number of bonds is preserved. In this step, the central ligand completes its coordination shell as shown in Fig. 11.20. This step is endothermic by 34 kJ/mol and requires to overcome a barrier of 45 kJ/mol.

In order to investigate the sequence of the bond-forming and bond-breaking

process, I investigated the two dimensional energy surface, spanned by the bond-lengths of the two Fe–N bonds, shown in Fig. 11.21. Each point corresponds to a state where both bond-lengths are fixed to the corresponding values, while all other degrees of freedom were fully relaxed.

The energy surface exhibits a third local minimum, namely *F4*. In *F4*, both bonds are broken simultaneously. Its energy lies 56 kJ/mol above that of *F2*, and 11 kJ above the direct transition state from *F2* to *F3*. Both transition states *TS2* and *TS3* lie energetically 62 kJ/mol above *F2*. Thus I conclude that this indirect path is unfavorable and the bond-breaking and bond formation occur in a concerted fashion, that is via *TS1* in Fig. 11.21.



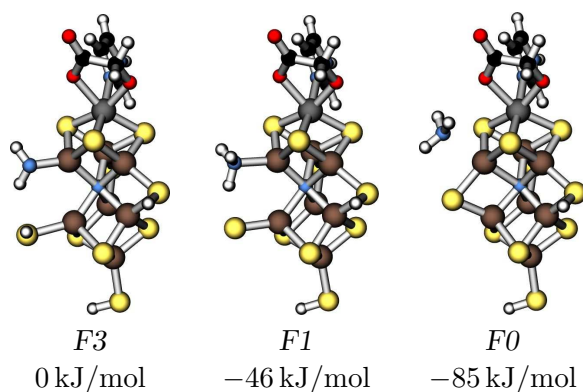
**Figure 11.21** Left: contour plot of the energy surface for a rearrangement after the fifth protonation, the transition from *F2* (upper left) to *F3* (bottom right).  $d(\text{N}-\text{Fe3})$  is the distance between Fe3 and N of the  $\text{NH}_2$  group.  $d(\text{N}_x-\text{Fe3})$  is the distance between Fe3 and the central ligand. The barrier for a direct transition from *F2* to *F3* via *TS1* is lower than the barrier for a transition via *TS2*, *F4*, and *TS3*. The energies are given relative to the starting state *F2*. The right-hand picture illustrates the energy surface with the corresponding structures positioned above the energy minima.

**Sixth protonation: Final ammonia detachment.** In the intermediate *F3*,  $\text{NH}_2$  exposes one lone pair which is readily protonated.

The most favorable pathway leads via an intramolecular proton transfer from the SH group. The proton transfer has a negligible barrier and is exothermic by 46 kJ/mol. The singly coordinated sulfur atom completes its coordination by restoring the sulfur bridge between the Fe sites. As this happens, ammonia is dissociated in a concerted fashion. Also this second reaction step has a negligible barrier and is exothermic, by 39 kJ/mol. It leads to *MH*, and thus closes the catalytic cycle.



As no external proton or electron transfer is required to proceed from *F2* to *F0*, the corresponding steps do not have to wait for any Fe-protein cycles to finish. All of them appear as one step in the Thorneley-Lowe scheme.



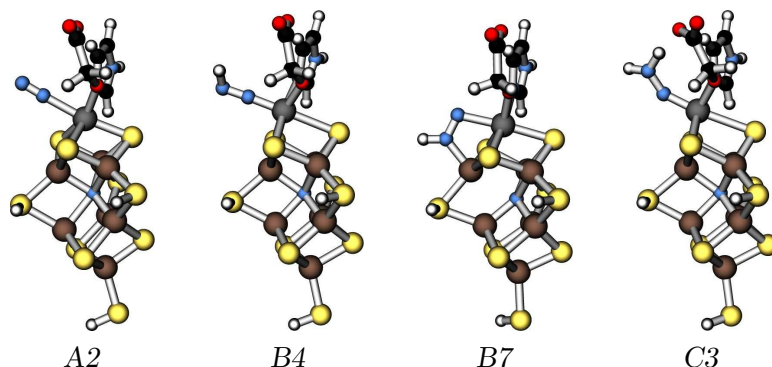
**Figure 11.22** Intermediates of the sixth protonation, the final ammonia detachment via intramolecular proton transfer. The energies are given relative to *F3*.

I considered also a side branch from *F3* to *MH2* proceeding via reduction of the cofactor in structure *F3* followed by an external proton transfer. Also this protonation from ammonium as proton donor proceeds easily. However, the resulting  $\text{NH}_3$  ligand is quite strongly bound to Fe. While dissociation is exothermic by 46 kJ/mol, it has a large barrier of 72 kJ/mol, the highest barrier in the whole process. Thus I conclude that this side branch is not relevant. Following dissociation of  $\text{NH}_3$ , the SH group reconnects to the second Fe, leading to *MH2*.

### 11.2.5 Molybdenum as $\text{N}_2$ reduction site

As shown in section 11.1.1 on page 134, the molybdenum atom is not a favorable nitrogen adsorption site: nitrogen bound to molybdenum is higher in energy by 50 kJ/mol than bound to iron. However, there have been various theoretical investigations [66–70, 99] and also experimental efforts [164, 172] to investigate this possible adsorption site. For this reason, I also discuss nitrogen coordination to molybdenum here. If I make the assumption that nitrogen binds to Mo, even though this scenario is at variance with the predictions presented in section 11.1.1 on page 134, my findings suggest a possible pathway for the first two critical protonations.

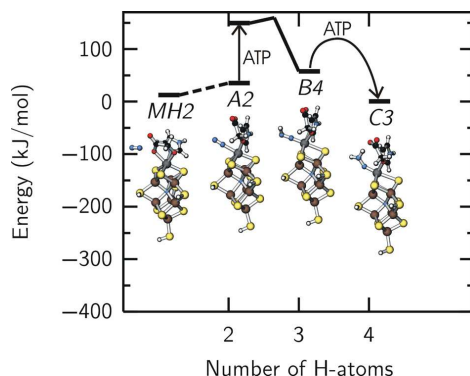
One freed coordination site at Mo, thus a monodentate homocitrate, is a prerequisite for  $\text{N}_2$  coordination to molybdenum. It is possible and energetically approximately neutral to dissociate one oxygen ligand of homocitrate from Mo, depending on the protonation state of Mo, as shown in section 10.1.1 on page 120.



**Figure 11.23** Intermediates of  $\text{N}_2$  bound to Mo and its first protonation products.

Nitrogen adsorption at penta-coordinate Mo (*A2*) is endothermic by 22 kJ/mol. This indicates that, even if the coordination site is vacant, dinitrogen binds to Mo only for fleetingly short periods of time. If protonation proceeds sufficiently easy, that the proton is transferred during these short periods, the dinitrogen may be stabilized bound to Mo. However, as seen in Fig. 11.24, protonation leading to *B4* is energetically unfavorable. *B4* is also 29 kJ/mol higher in energy than *B1*, with  $\text{N}_2\text{H}$  bound to Fe sites.

Durrant [69] proposed a transition of the protonated dinitrogen *B4* into a bridging position between Mo and Fe as in *B7* shown in Fig. 11.23. I found *B7* to be 79 kJ/mol higher in energy than *B4*.



**Figure 11.24** Energy scheme of the first steps in the conversion cycle starting from  $\text{N}_2$  bound to Mo. The barrier for Mo docking has not been calculated, which is indicated by a dashed line. See Figure 11.10 for details.

The second protonation step takes place at the terminal nitrogen atom *C3*. It was not possible to stabilize a structure with the second proton added to the

proximal N. The energy of *C3* is 47 kJ/mol higher than the energy of *C0*, the corresponding structure in the main reaction path.

Thus my calculations do neither support that (1) a coordination site at Mo becomes available for N<sub>2</sub> binding nor that (2) dinitrogen binds, if this coordination site is vacant. Thus I conclude that dinitrogen reduction will not take place at the molybdenum atom.

## 11.3 Discussion

So far, I presented a mechanism for nitrogen fixation with all intermediate states. Furthermore I discussed a wide range of side-branches, which had to be considered in order to identify the most favorable pathway.

The comparison of N<sub>2</sub>-binding and the critical first protonation step with previous studies with [104] or without [63, 69] central ligand determines the present mechanism as favorable: according to my calculations, nearly all intermediates of the mechanism proposed here are lower in energy than those of earlier proposals. The highest barrier between two intermediates is 66 kJ/mol, consistent with the experimental turn-over rate of the enzyme.

### 11.3.1 Required properties of the cofactor

While the central ligand does not significantly influence the protonation sites at the cofactor before N<sub>2</sub> binds, its influence becomes dominant as soon as N<sub>2</sub> binds. The main feature of the central ligand is its ability to form a variable number of bonds to the six Fe atoms. The central ligand changes its coordination from six-fold to five-fold and four-fold. This allows other ligands such as nitrogen and sulfur to form and cleave bonds to the Fe sites without deviations from the preferred tetrahedral coordination of the latter. This is particularly apparent when dinitrogen binds: while the Fe site to which dinitrogen binds in the axial binding mode *A0* maintains its coordination shell by giving up its sulfur bridge, the bridged binding mode *A1* would result in an unfavorable five-fold coordination of one Fe site, if the latter would not give up its bond to the central ligand.

A second important ingredient is the flexibility of the cofactor. I found states where two of the Fe sites are bridged by a sulfur bridge, by a dinitrogen bridge and by a single nitrogen atom. The cofactor allows the Fe sites to vary their distance and the cavity provides sufficient space for the opening of the SH bridge. While sulfur prefers a bond angle close to or even smaller than 90°, nitrogen prefers an angle closer to the tetrahedral angle of 109° as in *D0*. This change can be accommodated best if site Fe3 is pulled out of the cluster. This movement of Fe3 is made possible by cleaving its bond to the central ligand.

The first protonation of  $N_2$ , which is the critical reaction step, is facilitated by bonds between dinitrogen and two Fe sites. These bonds activate the triple bond by creating a bonding configuration already analogous to  $N_2H_2$ .

Another difficult reaction step is the cleavage of the N–N bond. This step requires the transfer of two electrons with antiparallel spins from the cofactor to dinitrogen. The delocalized electron system of the cofactor facilitates double electron transfer, because it results in only small changes of the average oxidation state. Secondly having the nitrogen next to Fe sites with antiparallel spins, the cofactor can easily offer a singlet electron pair.

Also the final reaction step, namely the dissociation of the second ammonia from  $F2$ , relies on the special features of the cofactor: it would have one of the largest barriers in the mechanism, if the sulfur bridge would not be restored. The dissociation of the ammonia is part of a substitution reaction, in which the coordination number of the Fe site is restored by the sulfur atom as soon as ammonia leaves.

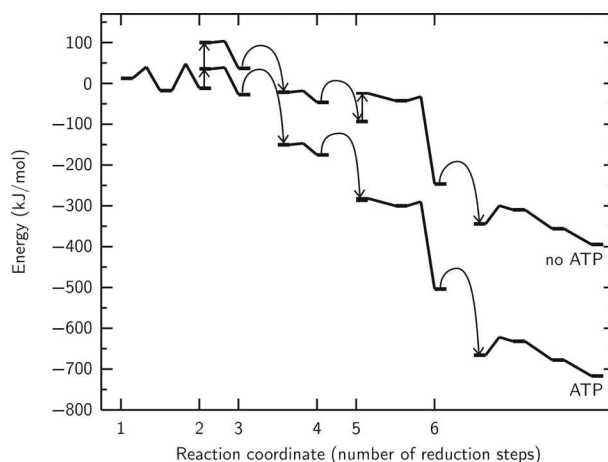
### 11.3.2 Binding site

The cofactor has an approximate three-fold symmetry, which is broken by the ligands and the protein environment. As long as the protein environment is not taken into account, as in the present study, the energetics will proceed similar for all three equivalent orientations. Nevertheless, the position of the proton channel in the protein indicates that nitrogen fixation proceeds near the iron sites 3 and 7, see section 7.1.3 on page 85. Furthermore the cavity in this region provides sufficient space to accommodate nitrogen bound to an Fe site, see section 11.1.4 on page 143.

My calculations do not allow to distinguish between Fe3 and Fe7 as potential docking sites of dinitrogen. I found that the axial binding mode *A0* is slightly more stable, that is by 14 kJ/mol, on Fe7 than on Fe3. The bridged mode *A1* accessed via the axial binding at Fe7 is more stable by 24 kJ/mol than the other variant. I expect that, due to the large motion of the sulfur atom, the opening of the SH bridge may be strongly affected by the shape and the specific interactions of the cavity. Thus, in this case, environment effects may be larger than the calculated energy differences, so that no conclusive answer regarding the initial binding site can be given at this point.

### 11.3.3 Dependence of the results on the choice of the hydrogen chemical potential

As pointed out in section 10.3, the energy profile depends on the choice of the chemical potential for protons and electrons, which are combined in  $\mu_H$ . Note



**Figure 11.25** Energy schemes of nitrogen conversion cycles at different chemical potentials of hydrogen ( $\mu_H$ ). The upper line assumes that no energy of MgATP is used for conversion while the lower line assumes all the energy gained from MgATP hydrolysis is used for the conversion.

that only the energies of those reaction steps, for which the number of hydrogen atoms change are affected by the choice of  $\mu_H$ . I have chosen a value for  $\mu_H$  based on a comparison of my calculations with experiment as described in section 10.3. Nevertheless, there is a large uncertainty on the chosen value. In order to relate my results to other calculations it is necessary to understand how the results depend on  $\mu_H$ .

The energy of each protonation step and thus the reaction energy of the entire nitrogen reduction cycle depends on the hydrogen chemical potential  $\mu_H$ . This value cannot be obtained from calculations as it depends on the chemical environment, the pH value, the reduction potential and other properties in the vicinity of the cluster.

A value for  $\mu_H$  may be obtained from the experimental reaction energy of the entire cycle. However, previous estimates differ substantially depending on the assumptions made in the analysis.

- Alberty [173] approximated the standard Gibbs energy for  $N_2$  conversion with  $\Delta_r G^\circ = -463.18$  kJ/mol for the chemical reaction defined via one specific set of educts and products (ferredoxin with a reduction potential of  $-0.403$ ,  $H^+$ ,  $NH_4^+$ ).
- For the biochemical equation defined for equilibrium concentrations of  $H_3O^+/H_2O$  and  $NH_4^+/NH_3$  at pH=7 with the same ferredoxin he obtained the transformed Gibbs energy  $\Delta_r G'^\circ = -63.62$  kJ/mol.

I rationalized the choice of  $\mu_H$  used in the present work, assuming  $\Delta_r E = 0$  for the first reduction and protonation of the resting state. This choice of  $\mu_H$  leads to  $\Delta_r E = -404 \text{ kJ/mol}$  for the whole reaction. The Gibbs free energy is related to the calculated reaction energy by  $\Delta_r G = \Delta_r U - T\Delta_r S + \Delta_r(pV)$  with  $\Delta_r U \approx \Delta_r E$ .

Most previous calculations chose the chemical potential  $\mu_H$  equal to one-half of the energy of a hydrogen molecule, which is  $36 \text{ kJ/mol}$  below my choice for  $\mu_H$  and leads to a reaction energy of  $\Delta_r E = -189 \text{ kJ/mol}$ .

Figure 11.25 shows the complete reaction path with two different chemical potentials of hydrogen. The profile labeled “no ATP” corresponds to the value for the chemical potential used in the present study. The profile labeled “ATP” corresponds to the maximum value of  $\mu_H$  consistent with the unprotonated resting state. Note that only the energy differences between states with different numbers of hydrogen atoms are affected.

### 11.3.4 The role of $\text{H}_2$

It is known from experiment that  $\text{H}_2$  is always produced by nitrogenase during  $\text{N}_2$  turnover [86]. Reduction equivalents, transferred to the MoFe-protein are lost in  $\text{H}_2$  production.  $\text{H}_2$  production can not be avoided by large partial pressures of  $\text{N}_2$  [84]. The amount of  $\text{H}_2$  production per  $\text{N}_2$  reduction in the limit of large  $\text{N}_2$  partial pressures is still under dispute. While Rivera-Ortiz and Burris [83] obtained 0.56–0.9  $\text{H}_2$  per  $\text{N}_2$ , Newton et al. [5] suggested a minimum stoichiometry of 1:1. Under ambient conditions larger amounts of  $\text{H}_2$  are produced.

My calculations indicate that  $\text{H}_2$  production competes with  $\text{N}_2$  reduction. Dihydrogen is always produced whenever a hydride is formed at an iron site. This happens if an excess of protons is transferred to the cofactor in the absence of bound  $\text{N}_2$ . If  $\text{N}_2$  is docked to the cofactor, these protons will reduce dinitrogen. Hence, our calculations are consistent with a production of non-stoichiometric amounts of  $\text{H}_2$ .

## 11.4 Conclusion

In the present study I made an attempt of a comprehensive investigation of possible pathways for nitrogen fixation at the Fe-Mo cofactor. The catalytic cycle that emerged from this study differs from previous proposals.

The distinct features of the mechanism are the following: Dinitrogen docks after two sulfur bridges of the cofactor have been protonated. Upon  $\text{N}_2$  binding, a sulfur bridge opens.  $\text{N}_2$  converts from an axial binding mode into a dinitrogen bridge, replacing the former sulfur bridge. After the second protonation,  $\text{N}_2\text{H}_2$  converts into a  $\pi$ -complex with one Fe-site, being further stabilized by a  $\sigma$ -bond to the

other Fe-site. The first ammonia is desorbed from  $\text{NH-NH}_3$  in a configuration, where NH bridges both Fe sites. The desorption requires a singlet electron pair, which is readily supplied by the two antiferromagnetically Fe atoms, to which dinitrogen is bound. One protonation further, the resulting  $\text{NH}_2$  loses one bond to an Fe-site and thus opens the bridge. Following an intramolecular protonation, the complex expels the second ammonia in a substitution reaction, which closes the sulfur bridge as the bond to ammonia cleaves. The central nitrogen ligand of the cofactor plays a critical role in the mechanism by balancing the coordination of the Fe-sites forming and breaking bonds to dinitrogen and its reduction products.

I provided a rationalization of the mechanistic features relevant to accomplish critical reaction steps, which may be useful to develop model systems for nitrogen fixation.





## 12 Nitrogenase and Acetylene

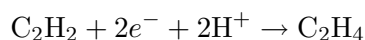
*An experiment is a question which science poses to Nature  
and a measurement is the recording of Nature's answer.*

Max Planck

### 12.1 Introduction

Nitrogenase is not only able to catalyze the conversion of  $\text{N}_2$  to  $\text{NH}_3$ , but can also reduce a number of other substrates. One of the most detailed studied among these alternative substrates is acetylene,  $\text{C}_2\text{H}_2$ .

While  $\text{N}_2$  is fully reduced to  $\text{NH}_3$  by the enzyme,  $\text{C}_2\text{H}_2$  is only reduced to  $\text{C}_2\text{H}_4$ , ethylene [174]. No further reduction to ethane,  $\text{C}_2\text{H}_6$ , takes place at the wild-type enzyme.



$\text{C}_2\text{D}_2$  as substrate allows to study the stereo-selectivity of the reduction. It is nearly quantitatively converted to *cis*- $\text{C}_2\text{D}_2\text{H}_2$ , only about 4% of *trans* product is found [175, 176].

The main reason why  $\text{C}_2\text{H}_2$  is better studied than  $\text{N}_2$  is the fact that acetylene binds to less reduced levels of the cofactor than  $\text{N}_2$  does. This makes it easier to access the  $\text{C}_2\text{H}_2$  binding mode experimentally. While dinitrogen is not able to bind to states less reduced than  $\text{E}2\text{H}2^1$  as discussed in section 11.1.5 on page 145, EPR/ENDOR experiments [177] show that  $\text{C}_2\text{H}_2$  even interacts with the resting state  $\text{E}0\text{H}0$  of the cofactor. Kinetic studies [178], however, conclude that  $\text{C}_2\text{H}_2$  is not reduced before bound to the  $\text{E}1\text{H}1$  state.

In the  $\text{N}_2$  conversion process,  $\text{H}_2$  is a necessary by-product [86], as described in section 3.2.1 on page 20.  $\text{H}_2$  production takes reduction equivalents from  $\text{N}_2$  reduction. In general,  $\text{H}_2$  is also produced during the conversion of acetylene. However, in this case the enzyme is theoretically able to completely suppress hydrogen production at infinite partial pressure of  $\text{C}_2\text{H}_2$  [83].

---

<sup>1</sup>The notation is explained in section 10.2 on page 123.  $\text{E}x\text{H}y$  refers to the cofactor reduced by  $x$  electrons and protonated  $y$  times.

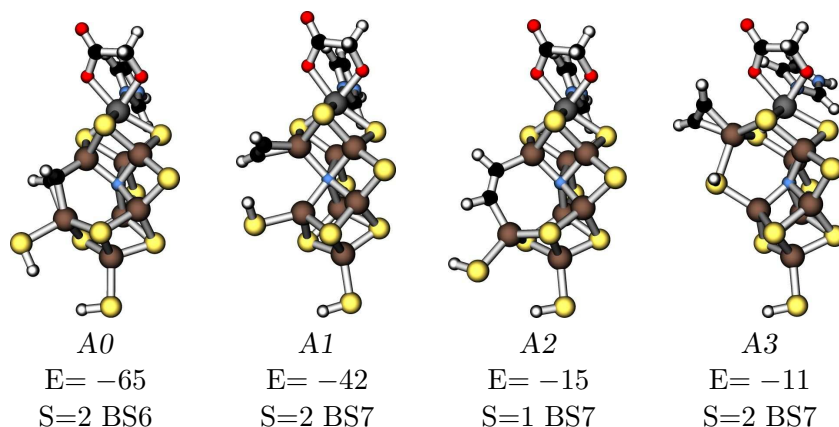
## 12.2 Computational Details

The cofactor of nitrogenase has been modeled as described in section 9.1.1. The same set of projector functions, the same plane-wave cutoff and the same unit-cell as in the calculations for  $N_2$  reduction have been used. In some cases, like in A2, even the starting structures have been obtained from the nitrogen conversion mechanism.

## 12.3 Results

### 12.3.1 Acetylene binding modes

I have investigated different possible binding modes at the E1H1 reduction level, suggested by experiment as the most oxidized level being able to bind and reduce  $C_2H_2$ . The stable and relevant binding modes are shown in Fig. 12.1. While dinitrogen can bind to Mo, although this mode is metastable,  $C_2H_2$  does not bind to Mo. Even after cleavage of one bond from Mo to homocitrate it was not possible to stabilize  $C_2H_2$  at the Mo site.  $C_2H_2$  drifted away.



**Figure 12.1** Binding modes of  $C_2H_2$  at the cofactor and their binding energy at the E1H1 level (kJ/mol) as well as their spin state. Negative energies indicate exothermic binding.

I found four stable binding modes in which  $C_2H_2$  is bound to one or more iron sites.

**A0:** The most stable mode exhibits  $C_2H_2$  bridging side-on between two iron sites forming a  $\pi$ -complex with both of them.  $C_2H_2$  replaces a sulfur bridge. One iron site is pulled out of the cage and its bond to the central ligand is broken. In accordance with my findings for dinitrogen binding in section 11.1.1, the

central ligand adds flexibility to the cluster by providing a varying number of bonds to its iron neighbors. All iron sites are in a distorted tetrahedral coordination and in their high-spin state.

Acetylene is already somewhat activated, which can be seen in Fig. 12.1 from the bending of the H-C-C-H unit and also from the increase of the C-C bond length from 1.21 Å in vacuum to 1.35 Å in *A0*.

This structure is, with a binding energy of 65 kJ/mol, the energetically most favorable of the binding modes which have been investigated.

- A1:**  $\eta^2$ -C<sub>2</sub>H<sub>2</sub> binding is energetically 17 kJ/mol higher than *A0*. The C-C bond is connected to one iron atom as a  $\pi$ -complex. Upon coordination of C<sub>2</sub>H<sub>2</sub> in this mode, the protonated sulfur bridge connected to the involved Fe site opens. The same lability of protonated sulfur bridges has been found upon N<sub>2</sub> binding to the cofactor in section 11.1.1. It preserves the tetrahedral environment of all iron atoms. They stay in their high-spin state. Even though the total spin is the same as in *A0*, the spin ordering is different: the spin of the iron site connected to the SH-group is flipped. In this binding mode, the C-C distance is 1.28 Å.
- A2:** In the  $\mu_2$  mode, C<sub>2</sub>H<sub>2</sub> bridges between the two iron atoms with each carbon connected to one iron site. The sp-hybridization, present in C<sub>2</sub>H<sub>2</sub> is converted in an sp<sup>2</sup>-hybridization. The C-C-H angles are about 117°. Two sp<sup>2</sup>-hybrid orbitals form the bonds to the iron atoms. The C-C bond is significantly enlarged from 1.21 Å in vacuum to 1.35 Å. However, the energy of this structure is rather high, the binding is only exothermic by 15 kJ/mol. A binding mode of this manner has been suggested from the experimentally found stereo-specificity of acetylene reduction [176]. However, as more stable binding modes exist, this one is not relevant for the conversion of C<sub>2</sub>H<sub>2</sub>.
- A3:** There is also an  $\eta^2$ -C<sub>2</sub>H<sub>2</sub> binding mode which preserves the sulfur bridge. In this case, the Fe site C<sub>2</sub>H<sub>2</sub> is bound to is pulled out of the cage and its bond to the central ligand is broken. C<sub>2</sub>H<sub>2</sub> binding is slightly exothermic with -11 kJ/mol. This mode is an intermediate during C<sub>2</sub>H<sub>2</sub> association to the cluster.

While the energies of the binding modes relative to each other can be expected to be represented rather accurately by my calculations, the absolute binding energies suffer from simplifications of the structural model. In the calculations, the zero for a binding energy is the energy of the cofactor on the one hand and isolated C<sub>2</sub>H<sub>2</sub> on the other hand. However, in the biological system, C<sub>2</sub>H<sub>2</sub> is provided by the protein and not obtained from vacuum. Moreover, the embedding energies of

the different binding modes in the protein have not been taken into account as they are not expected to vary much.

### 12.3.2 Barriers during acetylene binding

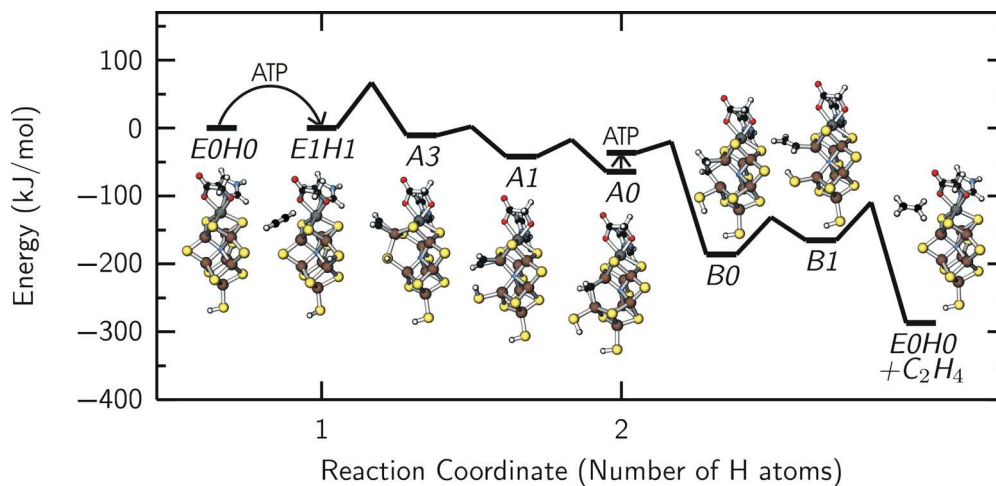
$C_2H_2$  coming from the protein environment can not directly access the most stable  $A0$  binding mode because major structural changes are required in the cofactor to accommodate that mode.

The route with the lowest barrier starts with  $C_2H_2$  binding to one Fe site, and pulling it out of the cluster, which results in  $A3$ . The barrier from free  $C_2H_2$  to  $A3$  is 67 kJ/mol. This is the largest barrier in the  $C_2H_2$  conversion process. It is in accordance with the experimental turnover rate. Next, the sulfur bridge opens by overcoming a barrier of 13 kJ/mol and  $A1$  is reached.

The two reactions of bridge-opening and  $C_2H_2$  association can—in principle—also occur in a concerted mechanism. However, in this case the barrier for reaching  $A1$  is 76 kJ/mol and thus substantially higher than the 67 kJ/mol via  $A3$ . Therefore,  $C_2H_2$  first associates and then the sulfur bridge opens.

The barrier for intercalation of  $C_2H_2$  between the two iron sites, which separates  $A1$  from  $A0$ , is less than 25 kJ/mol and can readily be overcome.

A schematic view of the energetics of the binding process as well as the following protonation is given in Fig. 12.2.



**Figure 12.2** Energy scheme of acetylene binding and reduction. The arrows indicate protonations. Their lengths depends on the choice of  $\mu_H$ . Intermediates connected by bold lines contain the same number of atoms. The energy difference between these is independent of  $\mu_H$ .

state	barrier	energy
E0H0		0
E1H1		0
A3	67	-11
A1	13	-42
A0	<25	-65
B0	16	-186
B1	55	-165
E0H0+C <sub>2</sub> H <sub>4</sub>	56	-287

**Table 12.1** Energetics of the acetylene conversion mechanism. The intermediates are given in the order of the mechanism. The barrier refers to the reaction leading to the respective intermediate. The energy is given relative to the resting state, free C<sub>2</sub>H<sub>2</sub> and using a  $\mu_H$  in order to keep the reduction an protonation of the resting state energetically neutral, as discussed in section 10.3. All energies are given in kJ/mol.

### 12.3.3 Protonation

The acetylene molecule in *A0* exhibits two lone-pairs, sp<sup>2</sup> hybrid orbitals. After reduction, one of these lone-pairs is protonated. They both do not point out of the cluster but into the direction of the faces of the cofactor spanned by 4 iron sites. Thus the proton donor has to approach the cofactor on one of these faces.

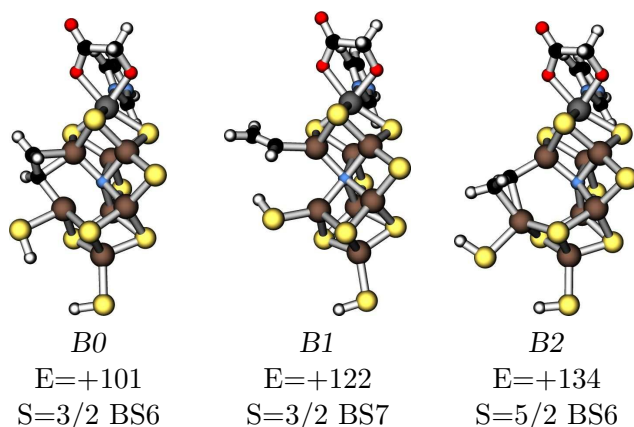
Using NH<sub>4</sub><sup>+</sup> as proton source as discussed in section 10.3 on page 128, the proton transfer is exothermic by 150 kJ/mol. Protonation results in cleavage of one  $\pi$ -complex bond, resulting in structure *B0* depicted in Fig. 12.3. The  $\pi$ -complex bond to the other iron atom remains. These structural rearrangements have a barrier of 16 kJ/mol. However, the barrier depends on the choice of the proton donor.

In a rearrangement, following the protonation, the C<sub>2</sub>H<sub>3</sub> converts into a usual  $\sigma$ -ligand bound to only one iron site, resulting in structure *B1* shown in Fig. 12.3. A reaction, in which the central ligand restores its sixth bond to its iron neighbors. This rearrangement is endothermic by 21 kJ/mol and has a barrier of 55 kJ/mol.

A third C<sub>2</sub>H<sub>3</sub> binding mode, *B2* has been identified, but it has a higher energy than those discussed previously and does not play any role in the reduction process.

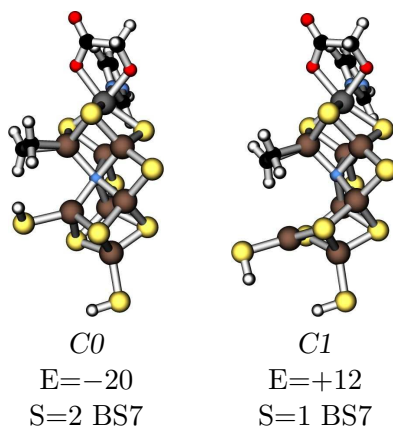
### 12.3.4 C<sub>2</sub>H<sub>4</sub> Production

In structure *B1*, the proximal CH group and the SH group are properly positioned for an intramolecular proton transfer. It is exothermic by 122 kJ/mol. The protonation of the C<sub>2</sub>H<sub>3</sub> fragment leads to C<sub>2</sub>H<sub>4</sub> which is immediately displaced by the closing sulfur bridge. The barrier for this process is 56 kJ/mol. It releases ethylene and restores the cofactor in its resting state.



**Figure 12.3** Intermediates after protonation of  $C_2H_2$  at the E1H1 reduction and protonation state. Energies are given in kJ/mol relative to  $C_2H_4$  cleavage. The total spin and the spin ordering is given.

This last internal proton transfer is the crucial step for the stereo-selectivity of the reduction process. In *B1*, the proton that has been added to  $C_2H_2$  is in the *cis*-position to the C–Fe bond. As this C–Fe bond is replaced by a C–H bond in the intermolecular proton transfer, the result is the *cis* product. Formation of the *trans* isomer would require a rotation around the C=C double bond of the  $C_2H_3$  fragment. This rotation has a barrier of 169 kJ/mol, which is impossible to overcome under reaction conditions. Therefore *cis*- $C_2D_2H_2$  is produced during  $C_2D_2$  conversion.



**Figure 12.4**  $C_2H_4$  bound to the cofactor and its binding energy in kJ/mol as well as the spin state.

If an intermolecular protonation is assumed,  $C_2H_4$  is not spontaneously dis-

placed but stays weakly bound to the cofactor. Two possible binding modes and their  $C_2H_4$  binding energies are shown in Fig. 12.4.

### 12.3.5 Acetylene binding energies at different reduction states of the cofactor

The more reduced the cofactor is, the greater is its affinity to  $C_2H_2$  as it can be seen in table 12.2. The energetic order of the different binding modes is preserved during reduction and protonation of the cluster. While binding is significantly more stable in the E2H2 state than in the E1H1 state discussed above, binding is even slightly endothermic in the resting state. The destabilization of A0, A1, and A2 in the resting state may be explained by the fact that the opening of the sulfur bridge leaves the sulfur atom coordinated to only one iron site. This is much more stable in case of a protonated sulfur bridge. A3 is the most stable binding mode in the resting state, as in this mode, the sulfur bridge is preserved.

	A0	A1	A2	A3
E0H0	+15			+9
E1H1	-65	-42	-15	-11
E2H2	-87	-58	-37	

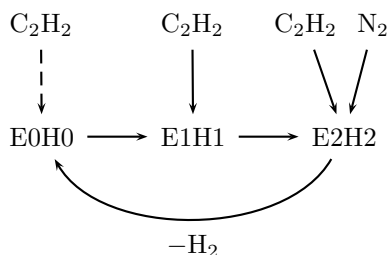
**Table 12.2**  $C_2H_2$  binding energies at different reduction- and protonation levels. Energies in kJ/mol, negative values indicate exothermic binding.

## 12.4 Discussion

**Inhibition:** It is found by experiment that dinitrogen is a weak competitive inhibitor of acetylene reduction but acetylene is an effective non-competitive inhibitor of dinitrogen reduction [83, 174, 179]. My calculations support an idea proposed by Davis et al. [180, 181]: acetylene binds to the cofactor at a state not sufficiently reduced for nitrogen to bind. Therefore it inhibits non-competitively because it lowers the pool for available  $N_2$  binding sites. Dinitrogen competitively inhibits acetylene reduction at the reduced state. As most acetylene is reduced in the oxidized state, the inhibition is weak.

Acetylene is able to bind and be reduced at the E1H1 level while dinitrogen needs at least the more reduced E2H2 level for being effectively bound and reduced, see section 11.1.5 on page 145. This is also illustrated in Fig. 12.5. Therefore most of the acetylene is bound and reduced at the E1H1 level and only a limited part of the cofactor molecules reach the E2H2 level. The EPR/ENDOR experiments [177] showing that acetylene already interacts with the resting state may be explained by

weak and reversible binding. My calculated binding energy of +9 kJ/mol indicates endothermic binding but does not rule out interaction.



**Figure 12.5** Scheme of binding of  $\text{C}_2\text{H}_2$  and  $\text{N}_2$  to FeMoco in the wild-type MoFe protein.  $\text{C}_2\text{H}_2$  weakly binds to the resting state E0H0 but is bound and reduced at the more reduced states E1H1 and E2H2.  $\text{N}_2$ , in contrast, binds at earliest at E2H2. Dihydrogen may be released from E2H2.

The statement that dinitrogen already binds at the E2H2 level should be understood in the sense that this is the reduction level of the cofactor. The MoFe-protein instead, is expected to be reduced by one more electron in that state [82]. Therefore the E2H2 reduction level of the cofactor corresponds to the E3H3 level of the protein expressed in the Thorneley-Lowe scheme.

**$\text{H}_2$  production:** Unlike  $\text{N}_2$ , acetylene is theoretically able to completely suppress the reduction of protons [83] at infinite pressure of  $\text{C}_2\text{H}_2$ . As pointed out in section 10.1.2 on page 122, nitrogenase produces  $\text{H}_2$  whenever a hydride is formed at an iron site. This hydride formation is energetically less favorable than protonation of a  $\mu^2$ -sulfur bridge. Therefore it can only occur next to an already protonated sulfur bridge, otherwise the proton is transferred to sulfur and not to iron. Thus  $\text{H}_2$  production is possible at earliest in the E2H2 state as no hydrogen atom is available at the resting state E0H0. This is also shown in Fig. 12.5.

While the state E2H2 is present at a high concentration in the absence of  $\text{C}_2\text{H}_2$ , when  $\text{N}_2$  is converted, it rarely occurs when acetylene is reduced. E2H2 is even less abundant if the pressure of  $\text{C}_2\text{H}_2$  is increased and therefore dihydrogen production may be suppressed by acetylene.

## Mutation

**His $\alpha$ 195:** Substitution of His $\alpha$ 195 by glutamine results in an MoFe-protein that does hardly reduce  $\text{N}_2$  but still reduces acetylene (and protons) at near wild-type rates [95, 97]. His $\alpha$ 195 provides a hydrogen bond to the  $\mu_2$ -sulfur bridge S2B and is the only proton source for that atom. This proton source is removed in the mutant strain. There are only two  $\mu_2$ -sulfur bridges which can be protonated,



namely S5A and S2B. Protonation of both of them is essential for the E2H2 state to be reached without hydride formation, as discussed above. Thus protonation of S2B is essential for N<sub>2</sub>-reduction, while it is not essential for acetylene binding as this readily occurs at the E1H1 level. Nor is it essential for H<sub>2</sub> production as in that case, a hydride is formed near S5A releasing H<sub>2</sub>.

In Glu $\alpha$ 195 nitrogenase, N<sub>2</sub> is not reduced but it still inhibits both proton and acetylene reduction. This has been interpreted by Christiansen et al. [3] “that acetylene, protons, and dinitrogen must occupy the same or closely overlapping binding sites within the MoFe protein.” Their interpretation is confirmed by my results of possible C<sub>2</sub>H<sub>2</sub> binding at the E2H2 level.

**Gly $\alpha$ 69:** Substitution of Gly $\alpha$ 69 by serine [92], cysteine, proline, glutamate, or asparaglate [3] results in an enzyme that is able to reduce N<sub>2</sub> at the normal rate but is hardly able to reduce acetylene [92]. In this mutant strains, acetylene was converted from a non-competitive to a competitive inhibitor of dinitrogen reduction.

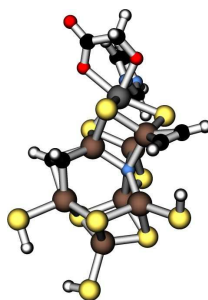
There is a structural interpretation of the effect of Gly $\alpha$ 69 mutation [3]: Gly $\alpha$ 69, as well as Val $\alpha$ 70 in the direct vicinity of the cofactor, are connected via a short  $\pi$ -helix to the P-cluster, reaching it with Cys $\alpha$ 62 as discussed in section 7.3.1 on page 92. The helix is assumed to communicate redox-dependent structural changes from the P-cluster to FeMoco. Movements of the chain may open a gate allowing substrates to bind to the cofactor. Even though it is not known how these residues might move upon reduction of the P-cluster, this movement is expected to be inhibited by mutation of the  $\alpha$ 69 residue inhibiting the binding of substrates already at the E1H1 level. Presumably, substrates are able to bind at the E2H2 level. This explains the competitive inhibition of dinitrogen reduction by acetylene. It may also be the case that the E1H1 level is the common binding state of both N<sub>2</sub> and C<sub>2</sub>H<sub>2</sub> in the  $\alpha$ 69-substituted protein.

**Arg $\alpha$ 96:** Mutation experiments of Arg $\alpha$ 96 may be explained in the same way: substitution of Arg $\alpha$ 96 by some smaller and non-positively charged amino acids creates proteins to which C<sub>2</sub>H<sub>2</sub> as well as CO can already bind in the resting state [94]. Arg $\alpha$ 96 is also connected to the P-cluster via a short chain as described in section 7.3.1 on page 92. Thus, in that case, a gate usually preventing effective binding of C<sub>2</sub>H<sub>2</sub> to the resting state may be opened by the mutation.

**Multiple binding sites:** Different EPR signals have been found during acetylene turnover in  $\alpha$ 195-Gln mutant [182]. This has been interpreted as two C<sub>2</sub>H<sub>2</sub> molecules simultaneously bound to the cofactor. Using the isolated cofactor, PhSH as coordination and europium amalgam cathode as reduction agent, it has also

been found that FeMoco-PhSH can simultaneously coordinate several substrate molecules to activate them for the subsequent reactions.

According to my calculations it is possible that two  $\text{C}_2\text{H}_2$  molecules are simultaneously bound to the cofactor. If one molecule is bound according to *A0*, there is the possibility for binding another one in an  $\eta^2$ -manner as in *A1* involving two different iron sites. The binding energy of the second  $\text{C}_2\text{H}_2$  molecule is  $-32\text{ kJ/mol}$  in the E2H2 level, indicating weak binding. The resulting structure is illustrated in Fig. 12.6.



**Figure 12.6** Two acetylene molecules simultaneously binding to the cofactor.

**Stereo-selectivity:** While previous reports [174] showed that *C. pasteurianum* produced exclusively *cis*- $\text{C}_2\text{H}_2\text{D}_2$  from  $\text{C}_2\text{D}_2$ , recent investigations [175, 176] reported that small amounts (4%) of the  $\text{C}_2\text{H}_2\text{D}_2$  product was the *trans* isomer. Production of mainly *cis*- $\text{C}_2\text{H}_2\text{D}_2$  is confirmed by my results. A production of the *trans* isomer would require to overcome a high barrier for a rotation around a double bond.

## 12.5 Conclusion

The catalytic conversion of  $\text{C}_2\text{H}_2$  to  $\text{C}_2\text{H}_4$  by nitrogenase offers a possibility to verify a proposed mechanism by comparison to a large amount of experimental data. As  $\text{C}_2\text{H}_2$  binds to less reduced forms of the cofactor than  $\text{N}_2$  does, the  $\text{C}_2\text{H}_2$  binding modes are easier to access experimentally.

I have proposed an acetylene conversion mechanism based on my first-principles calculations that is in general accordance with experiment. It explains the non-competitive inhibition of  $\text{N}_2$  conversion by  $\text{C}_2\text{H}_2$  as well as the weak competitive inhibition of  $\text{C}_2\text{H}_2$  conversion by  $\text{N}_2$ . It also accounts for the fact that  $\text{C}_2\text{H}_2$  can completely suppress  $\text{H}_2$  production of nitrogenase. Correspondingly the effect of mutations and the stereo-selectivity of the reaction are explained.

---

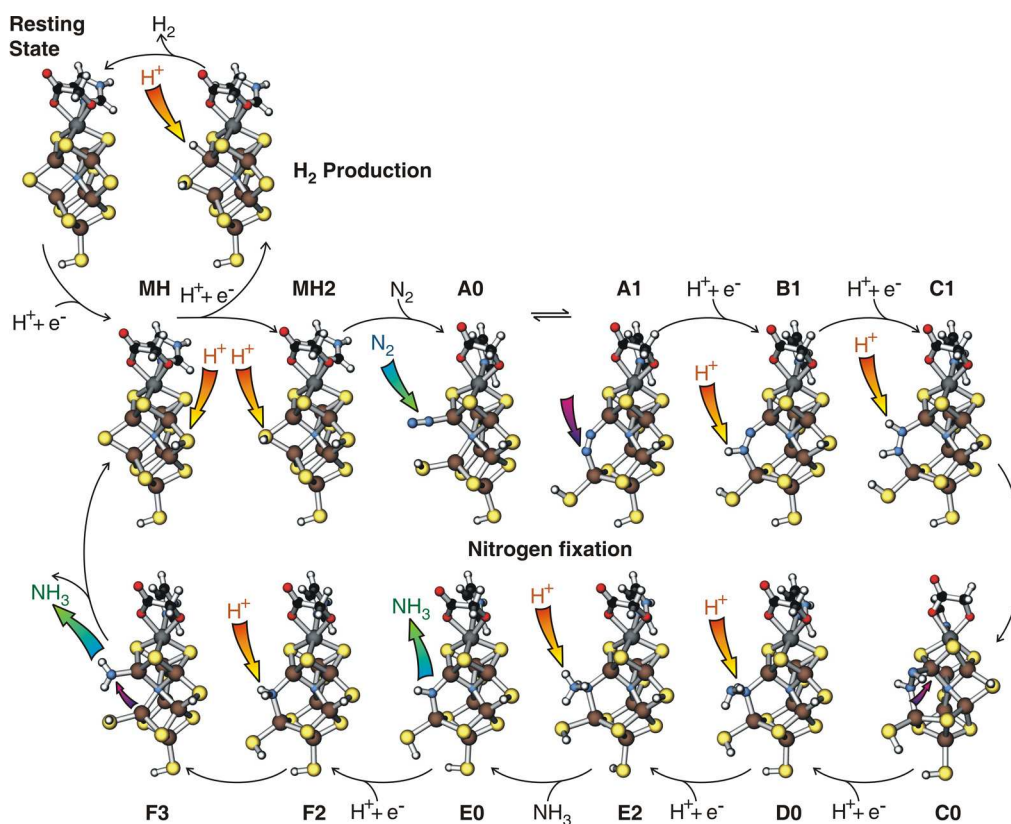
The general chemical reactivity of the cofactor with  $\text{C}_2\text{H}_2$  is similar to its reactivity with  $\text{N}_2$ . The general common features are that a sulfur bridge is destabilized by protonation and that the substrate is bound to multiple iron sites.

The good agreement of the proposed  $\text{C}_2\text{H}_2$  conversion supports the mechanism of  $\text{N}_2$  conversion, found by the same methodology.



## 13 Summary

In this work, I described the atomistic reaction mechanism of biological nitrogen fixation at the FeMo cofactor of nitrogenase, containing the recently detected central nitrogen ligand, as it follows from DFT calculations.



**Figure 13.1** Overview of structures and reactions of H<sub>2</sub> production and nitrogen fixation at the FeMo-cofactor of nitrogenase.

The resting charge state of the FeMo-cofactor of nitrogenase has been found to be  $[MoFe_7S_9N]^{\pm 0}$ .

Reductions and protonations of the cofactor are required prior to N<sub>2</sub> binding. An investigation of the protonation cascade results in a ping-pong mechanism for electron and proton transfer. Protons attach to the  $\mu_2$ -bridging sulfur atoms.

While one proton is provided to the cofactor by a proton path leading to S2B, shown on the right of each structure in Fig. 13.1, a series of protons is provided by a path leading to S5A, on the left side of each structure in Fig. 13.1. Double proton delivery by the same path leads to a side reaction:  $\text{H}_2$  production.

Dinitrogen binds to the cofactor after two or three of the sulfur bridges have been protonated (see structure *MH2* in Fig. 13.1). Dinitrogen binds axially to one of the Fe sites of the central cage. In order to maintain the tetrahedral coordination of the Fe atom, the sulfur bridge breaks open and converts into a non-bridging SH group on the other Fe-partner (*A0* in Fig. 13.1). The SH group remains isolated throughout the reaction until the very last step, where it plays a crucial role in the separation of the final reaction product. The axially bound nitrogen can insert, after overcoming a barrier of 66 kJ/mol, in between the two Fe sites, thus replacing the former SH bridge with a  $\text{N}_2$  bridge (*A1*). As the bridge is formed, the bond between one of these Fe-atoms and the central ligand breaks so that the tetrahedral coordination of the Fe atom is maintained. While the axial and the bridged configurations are separated by a considerable barrier, interconversion is possible on the time scale of the electron supply. Since they have similar energies, both states must be considered as potential protonation sites. I will later see that the two resulting branches of the reaction mechanism will join after the second proton transfer. In the following, I focus on the most favorable pathway, while the secondary branch will be discussed later.

The first proton is added to the dinitrogen in the bridged configuration. The barrier can readily be overcome. The resulting complex (*B1*) is not the most stable complex with this composition. The most favorable site to add the first proton is instead the central ligand with dinitrogen in the bridged binding mode. However, adding the proton to the central ligand involves a large barrier of 79 kJ/mol. This barrier corresponds to a rate on the order of one per second, which is lower than the turnover of the protein. Therefore, I consider the protonation of the central ligand as kinetically hindered.

The second proton is transferred to dinitrogen leading to (*C1*). The protonation is immediately followed, after overcoming a small barrier, by an intramolecular rearrangement: a bond shift rearranges the  $\text{N}_2\text{H}_2$  fragment so that it forms a  $\pi$ -complex (*C0*) with one Fe site, while the other Fe-atom stabilizes one of the lone pairs. A lone pair is a non-bonding orbital filled by two electrons. This intermediate state (*C0*) is common to all possible branches of the reaction, even for the one proceeding via the protonation of the central ligand.

$\text{N}_2\text{H}_2$  exposes one lone pair, which accepts the third proton. After protonation, the bond of the  $\pi$ -complex between Fe and  $\text{N}_2\text{H}_2$  converts into a single bond. In the resulting structure (*D0*) a single N atom bridges both Fe sites. Thus the bond to Fe has been shifted within the last two steps from one nitrogen atom to the other. The  $\text{NH}_2$  group coordinated to the bridging nitrogen atom exposes a lone

pair ready for the next protonation.

After protonation of the lone pair in (*E2*), the resulting ammonia molecule dissociates (*E0*) after overcoming a small barrier. In the gas phase this is the second critical step of the reaction. In order to break the bond between the nitrogen atoms, an electron pair must be supplied to the anti-bonding orbital so that two lone pairs result after dissociation. Thus the reaction is limited by the transfer rate from a potential electron source. The delocalized electron system of nitrogenase acts as an electron sponge, which supplies the electrons of high energy instantaneously with a fluctuation stretching the dinitrogen bond.

The NH-bridge exposes one lone pair, which is protonated (*F2*) with the next electron-transfer step. This in turn weakens the bond to the neighboring Fe sites. Only one of the participating Fe sites has an intact bond to the central ligand. Now, also the other Fe atom forms a bond to the central ligand, while giving up its bond to the NH<sub>2</sub> fragment (*F3*). After this step, the NH<sub>2</sub> group is bound to only one Fe site. This step is endothermic by 34 kJ/mol and has a barrier of 45 kJ/mol, the second largest barrier within the catalytic cycle.

The dissociation of the second ammonia requires another proton transfer. This transfer can occur internally from the SH group. It requires reorientation of the SH-group and the NH<sub>2</sub> group, which may require a rearrangement of the hydrogen bond network in the solvent. In my calculations the proton transfer proceeds with a negligible barrier and is exothermic by 46 kJ/mol. Once the proton is transferred, the sulfur atom inserts in between the two Fe sites, reestablishing the bridge present in the resting state. As a consequence, the ammonia molecule is displaced (*MH*). The bridge formation and the dissociation of the ammonia is exothermic by 39 kJ/mol.

A final proton transfer restores the initial configuration (*MH2*) with a protonated SH-group, that waits for the next nitrogen molecule to begin the next cycle.

Beside the above described most favorable path, there is one important side path that cannot be excluded. The first proton can be added to the axial dinitrogen binding mode (*A0*). In this case it is added to the terminal nitrogen atom and results in a configuration that is only 19 kJ/mol higher in energy than the protonated nitrogen bridge (*B1*) discussed earlier. The second proton is added to the proximal nitrogen atom. Following the second protonation the substrate rotates into the  $\pi$ -complex with one Fe site (*C0*) after overcoming a barrier of 44 kJ/mol, thus joining the reaction branch discussed above.

Molybdenum has also been investigated as dinitrogen binding site but can be excluded as N<sub>2</sub> binding to Mo is higher in energy by 50 kJ/mol than to Fe.

Binding and reduction of acetylene has been found to be possible already at a less reduced state than N<sub>2</sub> conversion. This explains the ability of C<sub>2</sub>H<sub>2</sub> to completely eliminate sacrificial H<sub>2</sub> production as well as the inhibition of N<sub>2</sub> conversion by C<sub>2</sub>H<sub>2</sub>. In analogy to the N<sub>2</sub> binding modes, destabilization of a sulfur bridge

and its replacement by the substrate is crucial for acetylene binding. After an intermolecular and an intramolecular protonation step, ethylene is released.

The mechanisms I proposed based on my state-of-the-art first-principles calculations differ substantially from the picture previously envisaged. The unique properties of the cofactor may lend guidance to the search for new biomimetic systems able to selectively activate and break strong covalent bonds.



## List of Tables

2.1	Structures of nitrogenase proteins determined by crystallographic analysis. . . . .	9
2.2	Hyperfine parameters of the resting state of FeMoco . . . . .	12
3.1	Standard reduction potentials of H <sub>2</sub> production and N <sub>2</sub> reduction. . . . .	21
3.2	Substitutions of amino acids in the vicinity of the cofactor . . . . .	22
5.1	Waiting times for overcoming a reaction barrier . . . . .	55
5.2	Spin contamination in various systems. . . . .	59
5.3	Comparison of collinear and non-collinear DFT calculations. . . . .	62
6.1	Magnetic hyperfine parameters of [Fe <sub>3</sub> S <sub>4</sub> ] <sup>+</sup> . . . . .	74
6.2	Calculated magnetic hyperfine parameters of [Fe <sub>3</sub> S <sub>4</sub> ] <sup>+</sup> . . . . .	76
7.1	Amino acids offering mobile proton transport groups . . . . .	84
7.2	The proton path from the cofactor to the surface. . . . .	87
8.1	Experimental atomization energies compared to my calculations . . . . .	96
8.2	Calculated data concerning N <sub>2</sub> and its reduction products . . . . .	96
9.1	Core configuration and number of projector functions and pairs of partial waves per angular-momentum state ( $\ell, m$ ) with s,p and d-character as used throughout the work for this thesis. . . . .	100
9.2	Total spin of different charge states of FeMoco . . . . .	102
9.3	Energies of the three sub-states of BS7. . . . .	108
9.4	The structure of the resting state of FeMoco compared to experiments. . . . .	114
9.5	Bond length with C, N, or O as central ligand compared to the X-ray structure from the 1M1N data base. . . . .	116
9.6	Mössbauer isomer shifts $\delta$ and electric field gradients $\Delta E_Q$ of the iron sites. . . . .	117
10.1	First protonation energies. . . . .	121

10.2	Comparison between the unprotonated resting state, the protonated resting state, and the X-ray structure. . . . .	125
10.3	Change of bond length upon reduction, measured by EXAFS and compared to my DFT calculations . . . . .	126
11.1	Dinitrogen binding modes . . . . .	138
11.2	Binding energies when metal semicore states are treated as valence states. . . . .	139
11.3	Comparison of the functionals PBE and RPBE for N <sub>2</sub> binding energies (kJ/mol). . . . .	140
11.4	Embedding energies of different N <sub>2</sub> binding modes. . . . .	144
11.5	Binding energy of N <sub>2</sub> (kJ/mol) at different oxidation- and protonation states of the cluster. . . . .	145
12.1	Energetics of the acetylene conversion mechanism. . . . .	171
12.2	C <sub>2</sub> H <sub>2</sub> binding energies at different reduction- and protonation levels	173

## List of Figures

2.1	Ribbon representations of the MoFe-protein and the Fe-protein. .	8
2.2	The FeMo-cofactor and the P-cluster. . . . .	9
2.3	Complex between the two nitrogenase proteins . . . . .	10
3.1	The Fe-protein cycle . . . . .	17
3.2	The kinetic reaction scheme proposed by Thorneley and Lowe. . .	19
3.3	Residues around FeMoco of which mutation experiments have been performed. . . . .	23
5.1	The PAW wave functions . . . . .	41
5.2	Top: projector functions of the chlorine atom for two s-type partial waves, middle: p-type, bottom: d-type. . . . .	45
5.3	Decomposition of a wave function into auxiliary wave function and the two one-center expansions . . . . .	46
6.1	Structure of the $[\text{Fe}_3\text{S}_4]^+$ complex. . . . .	73
6.2	Spin projection factors in $[\text{Fe}_3\text{S}_4]^+$ . . . . .	75
6.3	Angles between the spins of the $[\text{Fe}_3\text{S}_4]^+$ cluster. . . . .	77
6.4	Illustration of the spins of $[\text{Fe}_3\text{S}_4]^+$ as obtained from DFT calculations. . . . .	78
7.1	Nitrogenase FeMoco with structure and nomenclature of the 1M1N [19] PDB data base. . . . .	82
7.2	Mechanism of proton transport by hopping and rotation. . . . .	83
7.3	Schematic view of the algorithm deciding, which atoms are surface atoms. . . . .	85
7.4	Proton path from the surface to the sulfur sites surrounding Fe7. .	86
7.5	FeMoco in one half of the MoFe protein with the distance to the protein surface indicated. . . . .	89
7.6	Distance from the cofactor to the protein surface. . . . .	90
7.7	The network of hydrogen bonds around the homocitrate. . . . .	91
7.8	Helix from the P-cluster to Arg $\alpha$ 96. . . . .	92
7.9	Helix from the P-cluster to Val $\alpha$ 70. . . . .	93

8.1	Density of states and the shape of the corresponding orbitals of the N <sub>2</sub> molecule. The p- $\pi^*$ and the p- $\sigma^*$ states are unoccupied. . . . .	95
8.2	Calculated data concerning N <sub>2</sub> and its reduction products . . . . .	97
9.1	Change of the cluster size at different charges. . . . .	103
9.2	Fe-Fe bond length at different charge states. . . . .	104
9.3	Theoretical vs. experimental structure: stereoview of the superimposed structures from X-ray diffraction (black) and the theoretical resting state (red). . . . .	105
9.4	Overview of the electronic density of states of FeMoco in its resting state . . . . .	106
9.5	Spin ordering of the resting state. . . . .	107
9.6	All 10 spin orderings of FeMoco consistent with $S = 3/2$ . . . . .	108
9.7	Projected densities of states in the resting state: iron sites. . . . .	109
9.8	COOPs of Fe-Fe bonds . . . . .	110
9.9	Metal-metal bonds between iron atoms. . . . .	111
9.10	COOPs of Mo-Fe bonds . . . . .	111
9.11	Metal-metal bonds between Mo and Fe. . . . .	112
9.12	Density of states projected on the p-states of the central nitrogen ligand. . . . .	113
10.1	Different protonation sites of FeMoco. . . . .	120
10.2	Possible hydrogen bond between homocitrate and histidine. . . . .	121
10.3	Protonation and H <sub>2</sub> production. . . . .	123
10.4	Protonation energies for different charge states of the cofactor. . . . .	124
10.5	Distribution of the Fe-Fe bond length of the resting state. . . . .	126
10.6	Stereoview of the resting state structures of <i>K. pneumoniae</i> and <i>A. vinelandii</i> FeMoco. . . . .	127
10.7	Energy contributions to one protonation step. . . . .	130
11.1	Dinitrogen binding modes and the corresponding binding energy in energetic order. . . . .	137
11.2	The model proposed by Sellmann and its destabilization by the central ligand. . . . .	138
11.3	Energy barrier of opening of the sulfur bridge and axial coordination of N <sub>2</sub> to FeMoco. . . . .	140
11.4	The two spin states of the open axial structure. . . . .	141
11.5	Transition from the axial to the bridged binding mode. . . . .	142
11.6	$\pi$ -backdonation . . . . .	142
11.7	Environment of the axial N <sub>2</sub> binding mode. . . . .	144
11.8	Reaction scheme of the nitrogen conversion cycle. . . . .	146

---

11.9	Overview of the main path of the reaction mechanism. . . . .	147
11.10	Energy scheme of the total nitrogen conversion cycle. . . . .	148
11.11	Intermediates of branch 1. . . . .	149
11.12	Possible structures after the first protonation of N <sub>2</sub> bound to FeMoco. . . . .	150
11.13	Energy scheme of branch 2. . . . .	152
11.14	Intermediates of branch 2. . . . .	152
11.15	Energy scheme of branch 3. . . . .	153
11.16	Intermediates of branch 3. . . . .	154
11.17	First four intermediates of branch 4. . . . .	155
11.18	Intermediates of an alternative route leading to cleavage of the N–N bond. . . . .	156
11.19	Energy scheme of an alternative route leading to cleavage of the N–N bond. . . . .	157
11.20	Intermediates of a rearrangement after the fifth protonation. . . . .	157
11.21	Contour plot of the energy surface for a rearrangement after the fifth protonation. . . . .	158
11.22	Intermediates of the sixth protonation. . . . .	159
11.23	Intermediates of N <sub>2</sub> bound to Mo and its first protonation products. . . . .	160
11.24	Energy scheme of the first steps in the conversion cycle starting from N <sub>2</sub> bound to Mo. . . . .	161
11.25	Energy schemes of nitrogen conversion cycles at different chemical potentials of hydrogen ( $\mu_H$ ). . . . .	163
12.1	Binding modes of C <sub>2</sub> H <sub>2</sub> at the cofactor . . . . .	168
12.2	Energy scheme of acetylene binding and reduction. . . . .	170
12.3	Binding modes of C <sub>2</sub> H <sub>3</sub> to the cofactor . . . . .	172
12.4	Binding modes of C <sub>2</sub> H <sub>4</sub> to the cofactor . . . . .	172
12.5	Scheme of binding of C <sub>2</sub> H <sub>2</sub> and N <sub>2</sub> to FeMoco in the wild-type MoFe protein. . . . .	174
12.6	Two acetylene molecules simultaneously binding to the cofactor. . . . .	176
13.1	Overview of the main path of the reaction mechanism. . . . .	179



## Bibliography

- [1] G.J. Leigh, editor *Nitrogen fixation at the Millennium* Elsevier (2002).
- [2] J.B. Howard and D.C. Rees: Structural Basis of Biological Nitrogen Fixation. *Chem. Rev.* **96**, 2965 (1996).
- [3] J. Christiansen, D.R. Dean, and L.C. Seefeldt: Mechanistic Features of the Mo-Containing Nitrogenase. *Annu. Rev. Plant Physiol. Plant Mol. Biol.* **52**, 269 (2001).
- [4] R.H. Holm, P. Kennepohl, and E.I. Solomon: Structural and Functional Aspects of Metal Sites in Biology. *Chem. Rev.* **96**, 2239–2314 (1996).
- [5] B.K. Burges and D.J. Lowe: Mechanism of Molybdenum Nitrogenase. *Chem. Rev.* **96**, 2983–3011 (1996).
- [6] R.R. Eady: Structure-Function relationships of Alternative Nitrogenases. *Chem. Rev.* **96**, 3013–3030 (1996).
- [7] B.K. Burges: The iron-molybdenum Cofactor of Nitrogenase. *Chem. Rev.* **90**, 1377–1406 (1990).
- [8] B.E. Smith, M.C. Durrant, S.A. Fairhurst, C.A. Gormal, K.L.C. Grönberg, R.A. Henderson, S.K. Ibrahim, T. Le Gall, and C.J. Pickett: Exploring the reactivity of the isolated iron-molybdenum cofactor of nitrogenase. *Coord. Chem. Rev.* **185-186**, 669 (1999).
- [9] D.C. Rees and J.B. Howard: Nitrogenase: standing at the crossroads. *Curr. Opin. Chem. Biol.* **4**, 559 (2000).
- [10] D.C. Rees: Great metalloclusters in enzymology. *Annu. Rev. Biochem.* **71**, 221 (2002).
- [11] J.W. Peters, M.H.B. Stowell, S. Michael, S.M. Soltis, M.G. Finnegan, M.K. Johnson, and D.R. Rees: Redox-Dependent Structural Changes in the Nitrogenase P-Cluster. *Biochemistry* **36**, 1181–1187 (1997).

- 
- [12] M.M. Georgiadis, H. Komiya, P. Chakrabarti, D. Woo, J.J. Kornuc, and D.C. Rees: Crystallographic Structure of the Nitrogenase Iron Protein from *Azotobacter vinelandii*. *Science* **257**, 1653 (1992).
- [13] S.M. Mayer, W.G. Niehaus, and D.R. Dean: Reduction of short chain alkynes by a nitrogenase  $\alpha$ -70Ala-substituted MoFe protein. *J. Chem. Soc. Dalton Trans.* **5**, 802 (2002).
- [14] J. Kim and D.C. Rees: Crystallographic structure and functional implications of the nitrogenase molybdenum-iron protein from *Azotobacter vinelandii*. *Nature* **360**, 553 (1992).
- [15] J. Kim and D.C. Rees: Structural Models for the Metal Centers in the Nitrogenase Molybdenum-Iron Protein. *Science* **257**, 1667 (1992).
- [16] M.K.Chan, J. Kim, and D.C. Rees: The Nitrogenase FeMo-Cofactor and P-Cluster Pair: 2.2 Å Resolution Structures. *Science* **260**, 792 (1993).
- [17] J.L. Schlessman, D. Woo, L. Joshua-Tor, J.B. Howard, and D.C. Rees: Conformational variability in structures of the nitrogenase iron proteins from *Azotobacter vinelandii* and *Clostridium pasteurianum*. *J. Mol. Biol.* **280**, 669 (1998).
- [18] P. Strop, P.M. Takahara, H.-J. Chiu, H.C. Angove, B.K. Burgess, and D.C. Rees: Crystal Structure of the All-Ferrous  $[4\text{Fe-4S}]^0$  Form of the Nitrogenase Iron Protein from *Azotobacter vinelandii*. *Biochemistry* **40**, 651 (2001).
- [19] O. Einsle, F.A. Tezcan, S.L.A. Andrade, B. Schmid, M. Yoshida, J.B. Howard, and D.C. Rees: Nitrogenase MoFe-Protein at 1.16 Å Resolution: A Central Ligand in the FeMo-Cofactor. *Science* **297**, 1696 (2002).
- [20] H. Schindelin, C. Kisker, J.L. Schlessman, J.B. Howard, and D.C. Rees: Structure of ADP  $\text{AlF}_4^-$ -stabilized nitrogenase complex and its implications for signal transduction. *Nature* **387**, 370 (1997).
- [21] B. Schmid, O. Einsle, H.-J. Chiu, A. Willing, M. Yoshida, J.B. Howard, and D.C. Rees: Biochemical and Structural Characterization of the Cross-Linked Complex of Nitrogenase: Comparison to the ADP- $\text{AlF}_4^-$ -Stabilized Structure. *Biochemistry* **41**, 15557 (2002).
- [22] J. Kim, D. Woo, and D.C. Rees: X-ray crystal structure of the nitrogenase molybdenum-iron protein from *Clostridium pasteurianum* at 3.0 Å resolution. *Biochemistry* **32**, 7104 (1993).



- [23] S.M. Mayer, D.M. Lawson, C.A. Gormal, S.M. Roe, and B.E. Smith: New insights into structure-function relationships in nitrogenase: a 1.6 Å resolution X-ray crystallographic study of *Klebsiella pneumoniae* MoFe-protein. *J. Mol. Biol.* **292**, 871 (1999).
- [24] I. Dance: The consequences of an interstitial N atom in the FeMo cofactor of nitrogenase. *Chem. Commun.* **3**, 324 (2003).
- [25] T. Lovell, T. Liu, D.A. Case, and L. Noodleman: Structural, Spectroscopic, and Redox Consequences of a Central Ligand in the FeMoco of Nitrogenase: A Density Functional Theoretical Study. *J. Am. Chem. Soc.* **125**, 8377 (2003).
- [26] V. Vraijmasu, E. Münck, and E.L. Bominaar: Density Functional Study of the Electric Hyperfine Interactions and the Redox-Structural Correlations in the Cofactor of Nitrogenase. Analysis of General Trends in  $^{57}\text{Fe}$  Isomer Shifts. *Inorg. Chem.* **42**, 5974 (2003).
- [27] J. Schimpl, H.M. Petrilli, and P.E. Blöchl: Nitrogen Binding to the FeMo-Cofactor of Nitrogenase. *J. Am. Chem. Soc.* **125**, 15772 (2003).
- [28] E. Münck, H. Rhodes, W.H. Orme-Johnson, L.C. Davis, W.J. Brill, and V.K. Shah: Nitrogenase Mössbauer and EPR spectroscopy - MoFe protein component from *Azotobacter vinelandii*. *Biochim. Biophys. Acta* **400**, 32 (1975).
- [29] J.-M. Mouesca, L. Noodleman, and D.A. Case: Analysis of the  $^{57}\text{Fe}$  Hyperfine Coupling Constants and Spin States in Nitrogenase P-Clusters. *Inorg. Chem.* **33**, 4819 (1994).
- [30] S.J. Yoo, H.C. Angove, V. Papaefthymiou, B.K. Burgess, and E. Münck: Mössbauer Study of the MoFe Protein of Nitrogenase from *Azotobacter vinelandii* Using Selective  $^{57}\text{Fe}$  Enrichment of the M-Centers. *J. Am. Chem. Soc.* **122**, 4926 (2000).
- [31] A.E. True, M.J. Nelson, R.A. Venters, W.H. Orme-Johnson, and B.M. Hoffmann: Iron-57 hyperfine coupling tensors of the FeMo cluster in *Azotobacter vinelandii* MoFe protein: determination by polycrystalline ENDOR spectroscopy. *J. Am. Chem. Soc.* **110**, 1935 (1988).
- [32] R.A. Venters, M.J. Nelson, P.A. McLean, A.E. True, M.A. Levy, B.M. Hoffman, and W.H. Orme-Johnson: ENDOR of the resting state of nitrogenase molybdenum-iron proteins from *Azotobacter vinelandii*, *Klebsiella pneumoniae*, and *Clostridium pasteurianum*. Proton, iron-57, molybdenum-95, and sulfur-33 studies. *J. Am. Chem. Soc.* **108**, 3487 (1986).

- [33] R.R. Eady, B.E. Smith, Z.H.L. Abraham, F.E. Dodd, J.G. Grossmann, L.M. Murphy, R.W. Strange, and S.S. Hasnain: Difference EXAFS Studies of Nitrogenase and Nitrite Reductase Enzymes in the Nitrogen Cycle. *Journal de physique* **C2**, 611 (1997).
- [34] J. Christiansen, R.C. Tittsworth, B.J. Hales, and S.P. Cramer: Fe and Mo EXAFS of *Azotobacter vinelandii* Nitrogenase in Partially Oxidized and Singly Reduced Forms. *J. Am. Chem. Soc.* **117**, 10017 (1995).
- [35] H.I. Liu and A. Filippoudi et al.: EXAFS studies of FeMo-cofactor and MoFe protein: Direct evidence for the long-range Mo-Fe-Fe interaction and cyanide binding to the Mo in FeMo-cofactor. *J. Am. Chem. Soc.* **116**, 2418 (1994).
- [36] J. Chen, J. Christiansen, R.C. Tittsworth, B.J. Hales, S.J. George, D. Coucouvanis, and S.P. Cramer: Iron EXAFS of *Azotobacter vinelandii* nitrogenase molybdenum-iron and vanadium-iron proteins. *J. Am. Chem. Soc.* **115**, 5509 (1993).
- [37] I. Harvey, R.W. Strange, R. Schneider, C.A. Gormal, C.D. Garner, S.S. Hasnain, R.L. Richards, and B.E. Smith: X-ray absorption spectroscopic studies of the binding of ligands to FeMoco of nitrogenase from *Klebsiella pneumoniae*. *Inorg. Chim. Acta* **275**, 150 (1998).
- [38] P.A. McLean, V. Papaefthymiou, W.H. Orme-Johnson, and E. Münck: Isotopic Hybrids of Nitrogenase. Mössbauer Study of MoFe Protein with Selective 57-Fe Enrichment of the P-Clusters. *J. Biol. Chem.* **262**, 12900 (1987).
- [39] K.K. Surerus, M.P. Hendrich, P. Christie, D. Rottgardt, W.H. Orme-Johnson, and E. Münck: Mössbauer and Integer Spin EPR of the Oxidized P-Clusters of Nitrogenase:  $P^{ox}$  is a Non-Kramers System with a Nearly Degenerate Ground Doublet. *J. Am. Chem. Soc.* **114**, 8579 (1992).
- [40] J.H. Spee, A.F. Arendsen, H. Wassink, S.J. Marritt, W.R. Hagen, and H. Haaker: Redox properties and electron paramagnetic resonance spectroscopy of the transition state complex of *Azotobacter vinelandii* nitrogenase. *FEBS Lett.* **432**, 55 (1998).
- [41] P.A. Lindahl, E.P. Day, T.A. Kent, W.H. Orme-Johnson, and E. Münck: Mössbauer, EPR, and magnetization studies of the *Azotobacter vinelandii* Fe protein. *J. Biol. Chem.* **260**, 11160 (1985).
- [42] A.C. Nyborg, J.L. Johnson, A. Gunn, and G.D. Watt: Evidence for a Two-Electron Transfer Using the All-Ferrous Fe Protein during Nitrogenase Catalysis. *J. Biol. Chem.* **275**, 39307 (2000).

- [43] H.C. Angove, S.J. Yoo, E. Münck, and B.K. Burgess: An All-ferrous State of the Fe Protein of Nitrogenase. *J. Biol. Chem.* **273**, 26330 (1998).
- [44] H.C. Angove, S.J. Yoo, B.K. Burgess, and Eckard Münck: Mössbauer and EPR Evidence for an All-Ferrous  $\text{Fe}_4\text{S}_4$  Cluster with  $S = 4$  in the Fe Protein of Nitrogenase. *J. Am. Chem. Soc.* **119**, 8730 (1997).
- [45] S. Sen, R. Igarashi, A. Smith, M.K. Johnson, L.C. Seefeldt, and J.W. Peters: A Conformational Mimic of the MgATP-Bound “On State” of the Nitrogenase Iron Protein. *Biochemistry* **43**, 1787 (2004).
- [46] B. Hinnemann and J.K. Nørskov: Modeling a Central Ligand in the Nitrogenase FeMo Cofactor. *J. Am. Chem. Soc.* **125**, 1466 (2003).
- [47] J.P. Perdew: Density-functional approximation for the correlation energy of the inhomogeneous electron gas. *Phys. Rev. B* **33**, 8822 (1986).
- [48] B. Hammer, L.B. Hansen, and J.K. Nørskov: Improved adsorption energetics within density-functional theory using revised Perdew-Burke-Ernzerhof functionals. *Phys. Rev. B* **59**, 7413 (1999).
- [49] P.J. Stephens, F.J. Devlin, C.F. Chabalowski, and M.J. Frisch: Ab Initio Calculation of Vibrational Absorption and Circular Dichroism Spectra Using Density Functional Force Fields. *J. Phys. Chem.* **98**, 11623 (1994).
- [50] H. Deng and R. Hoffmann: Wie die  $\text{N}_2$  Aktivierung durch den FeMo-Cofaktor der Nitrogenase vonstatten gehen könnte. *Angew. Chem.* **105**, 1125 (1993).
- [51] W. Plass: Electronic structure of the iron-molybdenum and alternative cofactors of nitrogenases: a comparison and its consequences. *J. Mol. Struct. (Theochem.)* **315**, 53 (1994).
- [52] I.G. Dance: The Binding and Reduction of Dinitrogen at an  $\text{Fe}_4$  Face of the FeMo Cluster of Nitrogenase. *Aust. J. Chem.* **47**, 979 (1994).
- [53] I. Dance: Theoretical investigations of the mechanism of biological nitrogen fixation at the FeMo cluster site. *J. Biol. Inorg. Chem. (JBIC)* **1**, 581–586 (1996).
- [54] I. Dance: Calculated details of a mechanism for conversion of  $\text{N}_2$  to  $\text{NH}_3$  at the FeMo cluster of nitrogenase. *Chem. Commun.* **97**, 165 (1997).
- [55] I. Dance: Understanding structure and reactivity of new fundamental inorganic molecules: metal sulfides, metallocarbohedrenes, and nitrogenase. *Chem. Commun.* **98**, 523 (1998).

- [56] F.B.C. Machado and E.R. Davidson: N<sub>2</sub> activation by iron-sulfur complexes - a theoretical contribution. *Theor. Chim. Acta.* **92**, 315 (1995).
- [57] K.K. Stavrev and M.C. Zerner: A Theoretical Model for the Active Site of Nitrogenase. *Chem. Eur. J.* **2**, 83 (1996).
- [58] K.K. Stavrev and M.C. Zerner: On the reduced and oxidized forms of the FeMo- cofactor of *Azotobacter vinelandii* nitrogenase. *Theor. Chem. Acc.* **96**, 141 (1997).
- [59] K.K. Stavrev and M.C. Zerner: Studies on the Hydrogenation Steps of the Nitrogen Molecule at the *Azotobacter vinelandii* Nitrogenase Site. *Int. J. Quantum Chem.* **70**, 1159 (1998).
- [60] S.-J. Zhong and C.-W. Liu: Possible binding modes for dinitrogen activation by the FeMo-cofactor in nitrogenase. *Polyhedron* **16**, 653 (1997).
- [61] P.E.M. Siegbahn, J. Westerberg, M. Svensson, and R.H. Crabtree: Nitrogen Fixation by Nitrogenases: A Quantum Chemical Study. *J. Phys. Chem. B* **102**, 1615–1623 (1998).
- [62] T.H. Rod, B. Hammer, and J.K. Nørskov: Nitrogen Adsorption and Hydrogenation on a MoFe<sub>6</sub>S<sub>9</sub> Complex. *Phys. Rev. Lett.* **82**, 4054 (1999).
- [63] T.H. Rod and J.K. Nørskov: Modeling the Nitrogenase FeMo Cofactor. *J. Am. Chem. Soc.* **122**, 12751–12763 (2000).
- [64] T.H. Rod, A. Logadottir, and J.K. Nørskov: Ammonia synthesis at low temperatures. *J. Chem. Phys.* **112**, 5343–5347 (2000).
- [65] R.K. Szilagy, D.G. Musaev, and K. Morokuma: Theoretical studies of biological nitrogen fixation. Part II. Hydrogen bonded networks as possible reactant and product channels. *J. Mol. Struct. (Theochem.)* **506**, 131 (2000).
- [66] R.K. Szilagy, D.G. Musaev, and K. Morokuma: Theoretical Studies of Biological Nitrogen Fixation. I. Density Functional Modeling of the Mo-Site of the FeMo-Cofactor. *Inorg. Chem.* **40**, 766 (2001).
- [67] M.C. Durrant: A molybdenum-centered model for nitrogenase catalysis. *Inorg. Chem. Comm.* **4**, 60–62 (2001).
- [68] M.C. Durrant: Controlled protonation of iron-molybdenum cofactor by nitrogenase: a structural and theoretical analysis. *Biochem. J.* **355**, 569 (2001).

- [69] M.C. Durrant: An Atomic-Level Mechanism for Molybdenum Nitrogenase. Part 1. Reduction of Dinitrogen. *Biochemistry* **41**, 13934 (2002).
- [70] M.C. Durrant: An Atomic-Level Mechanism for Molybdenum Nitrogenase. Part 2. Proton Reduction, Inhibition of Dinitrogen Reduction by Dihydrogen, and the HD Formation Reaction. *Biochemistry* **41**, 13946 (2002).
- [71] F. Barrière, C.J. Pickett, and J. Talarmin: Extended Huückel calculations on functional and structural models of the FeMo-cofactor of nitrogenase. *Polyhedron* **20**, 27 (2001).
- [72] T. Lovell, J. Li, T. Liu, D.A. Case, and L. Noodleman: FeMo Cofactor of Nitrogenase: A Density Functional Study of States  $M^N$ ,  $M^{OX}$ ,  $M^R$ , and  $M^I$ . *J. Am. Chem. Soc.* **123**, 12392 (2001).
- [73] T. Lovell, J. Li, D.A. Case, and L. Noodleman: Binding Modes for the First Coupled Electron and Proton Addition to FeMoco of Nitrogenase. *J. Am. Chem. Soc.* **124**, 4546 (2002).
- [74] T. Lovell, J. Li, D.A. Case, and L. Noodleman: FeMo cofactor of nitrogenase: energetics and local interactions in the protein environment. *J. Biol. Inorg. Chem. (JBIC)* **7**, 735 (2002).
- [75] T. Lovell, R.A. Torres, W.-G. Han, T. Liu, D.A. Case, and L. Noodleman: Metal Substitution in the Active Site of Nitrogenase  $MFe_7S_9$  ( $M = Mo^{4+}$ ,  $V^{3+}$ ,  $Fe^{3+}$ ). *Inorg. Chem.* **41**, 5744 (2002).
- [76] H. Lee, B.J. Hales, and B.M. Hoffman: Metal-Ion Valencies of the FeMo Cofactor in CO-Inhibited and Resting State Nitrogenase by  $^{57}Fe$  Q-Band ENDOR. *J. Am. Chem. Soc.* **119**, 11395 (1997).
- [77] M.G. Duyvis, H. Wassink, and H. Haaker: Nitrogenase of *Azotobacter vinelandii*: Kinetic Analysis of the Fe Protein Redox Cycle. *Biochemistry* **37**, 17345 (1998).
- [78] R.N.F. Thorneley and D.J. Lowe: The mechanism of *Klebsiella pneumoniae* nitrogenase action. Pre-steady-state kinetics of an enzyme-bound intermediate in  $N_2$  reduction and of  $NH_3$  formation. *Biochem. J.* **224**, 887 (1984).
- [79] D.J. Lowe and R.N.F. Thorneley: The mechanism of *Klebsiella pneumoniae* nitrogenase action. The determination of rate constants required for the simulation of the kinetics of  $N_2$  reduction and  $H_2$  evolution. *Biochem. J.* **224**, 895 (1984).

- [80] R.N.F. Thorneley and D.J. Lowe: The mechanism of *Klebsiella pneumoniae* nitrogenase action. Simulation of the dependences of H<sub>2</sub> evolution rate on component-protein concentration and ratio and sodium dithionite concentration. *Biochem. J.* **224**, 903 (1984).
- [81] K. Fisher, D.J. Lowe, and R.N. Thorneley: *Klebsiella pneumoniae* nitrogenase. The pre-steady-state kinetics of MoFe-protein reduction and hydrogen evolution under conditions of limiting electron flux show that the rates of association with the Fe-protein and electron transfer are independent of the oxidation level of the MoFe-protein. *Biochem. J.* **279**, 81 (1991).
- [82] K. Fisher, W. Newton, and D.J. Lowe: Electron Paramagnetic Resonance Analysis of Different *Azotobacter vinelandii* Nitrogenase MoFe-Protein Conformations Generated during Enzyme Turnover: Evidence for  $S = 3/2$  Spin States from Reduced MoFe-Protein. *Biochemistry* **40**, 3333 (2001).
- [83] J.M. Rivera-Ortiz and R.H. Burris: Interactions Among Substrates and Inhibitors of Nitrogenase. *J. Bacteriol.* **123**, 537 (1975).
- [84] F.B. Simpson and R.H. Burris: A Nitrogen Pressure of 50 Atmospheres Does Not Prevent Evolution of Hydrogen by Nitrogenase. *Science* **224**, 1095 (1984).
- [85] H.-I. Lee, K. S. Thrasher, D. R. Dean, W. E. Newton, and B. M. Hoffman: <sup>14</sup>N Electron Spin-Echo Envelope Modulation of the  $S = 3/2$  Spin System of the *Azotobacter vinelandii* Nitrogenase Iron-Molybdenum Cofactor. *Biochemistry* **37**, 13370 (1998).
- [86] K.L. Hadfield and W.A. Bulen: Adenosine triphosphate requirement of nitrogenase from *Azotobacter vinelandii*. *Biochemistry* **8**, 5103 (1969).
- [87] J.H. Guth and R.H. Burris: Inhibition of nitrogenase-catalyzed ammonia formation by hydrogen. *Biochemistry* **22**, 5111 (1983).
- [88] W.A. Bulen, R.C. Burns, and J.R. LeComte: Nitrogen fixation - hydrosulfite as electron donor with cell-free preparations of *Azotobacter vinelandii* and *Rhodospirillum rubrum*. *Proc. Natl. Acad. Sci. U.S.A.* **53**, 532 (1965).
- [89] G.E. Hoch, K.C. Schneider, and R.H. Burris: Hydrogen evolution and exchange, and conversion of N<sub>2</sub>O to N<sub>2</sub> by soybean root nodules. *Biochim. Biophys. Acta A* **37**, 273 (1960).
- [90] B.B. Jensen and R.H. Burris: Effect of high pN<sub>2</sub> and high pD<sub>2</sub> on ammonia production, hydrogen evolution, and hydrogen deuteride formation by nitrogenases. *Biochemistry* **24**, 1141 (1985).

- [91] D. Voet, J.G. Voet, and C.W. Pratt: *Lehrbuch der Biochemie* John Wiley & Sons (2002).
- [92] J. Christiansen, V.L. Cash, L.C. Seefeldt, and D.R. Dean: Isolation and Characterization of an Acetylene-resistant Nitrogenase. *J. Biol. Chem.* **275**, 11459 (2000).
- [93] P.M.C. Benton, M. Laryukhin, S.M. Mayer, B.M. Hoffman, D.R. Dean, and L.C. Seefeldt: Localization of a Substrate Binding Site on the FeMo-Cofactor in Nitrogenase: Trapping Propargyl Alcohol with an  $\alpha$ -70-Substituted MoFe Protein. *Biochemistry* **42**, 9102–9109 (2003).
- [94] P.M.C. Benton, S.M. Mayer, J.L. Shao, B.M. Hoffman, D.R. Dean, and L.C. Seefeldt: Interaction of Acetylene and Cyanide with the Resting State of Nitrogenase  $\alpha$ -96-Substituted MoFe Proteins. *Biochemistry* **40**, 13816 (2001).
- [95] M. Sørle, J. Christiansen, B.J. Lemon, J.W. Peters, D.R. Dean, and B.J. Hales: Mechanistic Features and Structure of the Nitrogenase  $\alpha$ -Gln195 MoFe Protein. *Biochemistry* **40**, 1540 (2001).
- [96] M.J. Dilworth, K. Fisher, C.-H. Kim, and W.E. Newton: Effects on Substrate Reduction of Substitution of Histidine-195 by Glutamine in the  $\alpha$ -Subunit of the MoFe Protein of *Azotobacter vinelandii* Nitrogenase. *Biochemistry* **37**, 17495 (1998).
- [97] C.H. Kim, W.E. Newton, and D.R. Dean: Role of the MoFe Protein  $\alpha$ -Subunit Histidine-195 Residue in FeMo-cofactor Binding and Nitrogenase Catalysis. *Biochemistry* **34**, 2798 (1995).
- [98] J. Shen, D.R. Dean, and W.E. Newton: Evidence for Multiple Substrate-Reduction Sites and Distinct Inhibitor-Binding Sites from an Altered *Azotobacter vinelandii* Nitrogenase MoFe Protein. *Biochemistry* **36**, 4884 (1997).
- [99] K.L.C. Grönberg, C.A. Gormal, M.C. Durrant, B.E. Smith, and R.A. Henderson: Why R-Homocitrate Is Essential to the Reactivity of FeMo-Cofactor of Nitrogenase: Studies on NifV<sup>-</sup>-Extracted FeMo-Cofactor. *J. Am. Chem. Soc.* **120**, 10613 (1998).
- [100] S.M. Mayer, C.A. Gormal, B.E. Smith, and D.M. Lawson: Crystallographic Analysis of the MoFe Protein of Nitrogenase from a nifV Mutant of *Klebsiella pneumoniae* Identifies Citrate as a Ligand to the Molybdenum of Iron Molybdenum Cofactor (FeMoco). *J. Biol. Chem.* **277**, 35263 (2002).

- [101] K. Fisher, M.J. Dilworth, and W.E. Newton: Differential Effects on N<sub>2</sub> Binding and Reduction, HD Formation, and Azide Reduction with  $\alpha$ -195His- and  $\alpha$ -191Gln-Substituted MoFe Proteins of *Azotobacter vinelandii* Nitrogenase. *Biochemistry* **39**, 15570 (2000).
- [102] H. Lee, M. Sørlie, J. Christiansen, R. Song, D.R. Dean, B.J. Hales, and B.M. Hoffman: Characterization of an Intermediate in the Reduction of Acetylene by the Nitrogenase  $\alpha$ -Gln<sup>195</sup> MoFe Protein by Q-band EPR and <sup>13</sup>C, <sup>1</sup>H ENDOR. *J. Am. Chem. Soc.* **122**, 5582 (2000).
- [103] S. Siemann, K. Schneider, M. Oley, and Müller: Characterization of a Tungsten-Substituted Nitrogenase Isolated from *Rhodobacter capsulatus*. *Biochemistry* **42**, 3846 (2003).
- [104] B. Hinnemann and J.K. Nørskov: Chemical Activity of the Nitrogenase FeMo Cofactor with a Central Nitrogen Ligand: Density Functional Study. *J. Am. Chem. Soc.* **126**, 3920 (2004).
- [105] U. Huniar, R. Ahlrichs, and D. Coucouvanis: Density Functional Theory Calculations and Exploration of a Possible Mechanism of N<sub>2</sub> Reduction by Nitrogenase. *J. Am. Chem. Soc.* **126**, 2588 (2004).
- [106] H.I. Lee, P.M.C. Benton, M. Laryukhin, R.Y. Igarashi, D.R. Dean, L.C. Seefeldt, and B.M. Hoffman: The Interstitial Atom of the Nitrogenase FeMo-Cofactor: ENDOR and ESEEM Show It Is Not an Exchangeable Nitrogen. *J. Am. Chem. Soc.* **125**, 5604 (2003).
- [107] Z. Cui, A.J. Dunford, M.C. Durrant, R.A. Henderson, and B.E. Smith: Binding Sites of Nitrogenase: Kinetic and Theoretical Studies of Cyanide Binding to Extracted FeMo-Cofactor Derivatives. *Inorg. Chem.* **42**, 6252 (2003).
- [108] C.J. Pickett: The Chatt cycle and the mechanism of enzymic reduction of molecular nitrogen. *J. Biol. Inorg. Chem.* **1**, 601 (1996).
- [109] T.H. Rod: *Density Functional Calculations and Modeling of the Biological Ammonia Synthesis* PhD thesis Technical University of Denmark (2000).
- [110] F. Bozso, G. Ertl, M. Grunze, and M. Weiss: Interaction of nitrogen with iron surfaces Fe(100) and Fe(111). *J. Catal.* **49**, 18 (1977).
- [111] I. Alstrup, I. Chorkendorff, and S. Ullmann: The dissociative chemisorption of nitrogen on iron(111) at elevated pressures. *Z. Phys. Chem.* **198**, 123 (1997).



- 
- [112] S. Dahl, A. Logadottir, R.C. Egeberg, J.H. Larsen, I. Chorkendorff, E. Törnqvist, and J. K. Nørskov: Role of Steps in N<sub>2</sub> Activation on Ru(0001). *Phys. Rev. Lett.* **83**, 1814 (1999).
- [113] P. Hohenberg and W. Kohn: Inhomogeneous electron gas. *Phys. Rev.* **136**, B864 (1964).
- [114] R.G. Parr and W. Yang: *Density-Functional Theory of Atoms and Molecules* Oxford Science Publishers (1989).
- [115] W. Kohn and L.J. Sham: Self-consistent equations including exchange and correlation effects. *Phys. Rev.* **140**, A1133 (1965).
- [116] R.M. Dreizler and E.K.U. Gross: *Density Functional Theory* Springer-Verlag (1990).
- [117] J.P. Perdew and Y. Wang: Accurate and simple analytic representation of the electron-gas correlation energy. *Phys. Rev. B* **45**, 13244 (1992).
- [118] A.D. Becke: Density-functional thermochemistry. I. The effect of the exchange-only gradient correction. *J. Chem. Phys.* **96**, 2155 (1992).
- [119] A.D. Becke: Density-functional thermochemistry. II. The effect of the Perdew-Wang generalized-gradient correlation correction. *J. Chem. Phys.* **97**, 9173 (1992).
- [120] A.D. Becke: Density-functional thermochemistry. II. The effect of the Perdew-Wang generalized-gradient correlation correction. *J. Chem. Phys.* **97**, 9173 (1992).
- [121] J.P. Perdew, K. Burke, and M. Ernzerhof: Generalized Gradient Approximation Made Simple. *Phys. Rev. Lett.* **77**, 3865 (1996).
- [122] P.E. Blöchl, C.J. Först, and J. Schimpl: Projector augmented wave method: ab initio molecular dynamics with full wave functions. *Bull. Mater. Sci.* **26**, 33 (2003).
- [123] R. Catlow, H. Shercliff, and S. Yip, editors *Handbook of Materials Modeling* chapter "Electronic structure methods: Augmented Waves, Pseudopotentials and the Projector Augmented Wave Method" by P. E. Blöchl, J. Kästner and C. J. Först Kluwer Academic Publishers (2004).
- [124] P.E. Blöchl: Projector augmented-wave method. *Phys. Rev. B* **50**, 17953 (1994).

- 
- [125] G. Kresse and D. Joubert: From ultrasoft pseudopotentials to the projector augmented-wave method. *Phys. Rev. B* **59**, 1758 (1999).
- [126] S.C. Watson and E.A. Carter: Spin-dependent pseudopotentials. *Phys. Rev. B* **58**, R13309 (1998).
- [127] C.G. Van de Walle and P.E. Blöchl: First-principles calculations of hyperfine parameters. *Phys. Rev. B* **47**, 4244 (1993).
- [128] D. Vanderbilt: Soft self-consistent pseudopotentials in a generalized eigenvalue formalism. *Phys. Rev. B* **41**, 7892 (1990).
- [129] L. Verlet: Computer "Experiments" on Classical Fluids. I. Thermodynamical Properties of Lennard-Jones Molecules. *Phys. Rev.* **159**, 98 (1967).
- [130] R. Car and M. Parrinello: Unified Approach for Molecular Dynamics and Density-Functional Theory. *Phys. Rev. Lett.* **55**, 2471 (1985).
- [131] J. Neugebauer: Algorithms for total energy minimisation. Talk at the "Workshop on Application of Density-Functional Theory in Condensed-Matter Physics, Surface Physics, Chemistry, Engineering, and Biology", Berlin (2003).
- [132] J. Neugebauer and C.G. Van de Walle: . *MRS Symp. Proc.* **408**, 43 (1996).
- [133] S. Ismail-Beigi and T. A. Arias: New algebraic formulation of density functional calculation. *Comp. Phys. Comm.* **128**, 1 (2000).
- [134] M. P. Teter, M. C. Payne, and D. C. Allan: Solution of Schrödinger's equation for large systems. *Phys. Rev. B* **40**, 12255 (1989).
- [135] G. Kresse and J. Furthmüller: Efficient iterative schemes for ab initio total-energy calculations using a plane-wave basis set. *Phys. Rev. B* **54**, 11169 (1996).
- [136] G.E. Uhlenbeck and S. Goudsmit: Spinning Electrons and the Structure of Spectra. *Nature* **117**, 264 (1925).
- [137] L.M. Sandratskii and P.G. Guletskii: Symmetrised method for the calculation of the band structure of non-collinear magnets. *J. Phys. F* **16**, L43 (1986).
- [138] J. Kübler, K.-H. Hock, J. Sticht, and A.R. Williams: Density functional theory of non-collinear magnetism. *J. Phys. F* **18**, 469 (1988).

- [139] T. Oda, A. Pasquarello, and R. Car: Fully Unconstrained Approach to Non-collinear Magnetism: Application to Small Fe Clusters. *Phys. Rev. Lett.* **80**, 3622 (1998).
- [140] D. Hobbs, G. Kresse, and J. Hafner: Fully unconstrained noncollinear magnetism within the projector augmented-wave method. *Phys. Rev. B* **62**, 11556 (2000).
- [141] C. Zener: Interaction between the d-Shells in the Transition Metals. II. Ferromagnetic Compounds of Manganese with Perovskite Structure. *Phys. Rev.* **82**, 403 (1951).
- [142] P.W. Anderson and H. Hasegawa: Considerations on Double Exchange. *Phys. Rev.* **100**, 675 (1955).
- [143] P. Kurz, G. Bihlmayer, S. Blügel, K. Hirai, and T. Asada: Comment on "Ultrathin Mn films on Cu(111) substrates: Frustrated antiferromagnetic order". *Phys. Rev. B* **63**, 96401 (2001).
- [144] I. Turek, J. Kudrnovsky, M. Divis, P. Franek, G. Bihlmayer, and S. Blügel: First-principles study of the electronic structure and exchange interactions in bcc europium. *Phys. Rev. B* **68**, 224431 (2003).
- [145] M.R. Antonio, B.A. Averill, I. Moura, J.J. Moura, W.H. Orme-Johnson, B.K. Teo, and A.V. Xavier: Core dimensions in the 3Fe cluster of Desulfovibrio gigas ferredoxin II by extended X-ray absorption fine structure spectroscopy. *J. Biol. Chem.* **257**, 6646 (1982).
- [146] T.A. Kent, B.H. Huynh, and E. Münck: Iron-sulfur proteins: Spin-coupling model for three-iron clusters. *Proc. Natl. Acad. Sci. U.S.A.* **77**, 6574 (1980).
- [147] R.H. Holm: Trinuclear cuboidal and heterometallic cubane-type iron-sulfur clusters - new structural and reactivity themes in chemistry and biology. *Adv. Inorg. Chem.* **38**, 1 (1992).
- [148] C. Krebs, T.F. Henshaw, J. Cheek, B.H. Huynh, and J.B. Broderick: Conversion of 3Fe-4S to 4Fe-4S Clusters in Native Pyruvate Formate-Lyase Activating Enzyme: Mössbauer Characterization and Implications for Mechanism. *J. Am. Chem. Soc.* **122**, 12497 (2000).
- [149] Y. Sanakis, A.L. Macedo, I. Moura, J.J.G. Moura, V. Papaefthymiou, and E. Münck: Evidence for Antisymmetric Exchange in Cuboidal  $[3\text{Fe-4S}]^+$  Clusters. *J. Am. Chem. Soc.* **122**, 11855 (2000).

- [150] J. Telser, H. Huang, H. Lee, M.W.W. Adams, and B.M. Hoffman: Site Valencies and Spin Coupling in the 3Fe and 4Fe ( $S = 1/2$ ) Clusters of *Pyrococcus furiosus* Ferredoxin by  $^{57}\text{Fe}$  ENDOR. *J. Am. Chem. Soc.* **120**, 861 (1998).
- [151] C. Fan, A.L.P. Houseman, P. Doan, and B.M. Hoffman: Conformational distribution in protein-bound  $[\text{3Fe-4S}]^+$  clusters: CW and pulsed EPR and iron-57 ENDOR of D. gigas hydrogenase. *J. Phys. Chem.* **97**, 3017 (1993).
- [152] Z. Hu, D. Jollie, B.K. Burgess, P.J. Stephens, and E. Münck: Moessbauer and EPR Studies of *Azotobacter vinelandii* Ferredoxin I. *Biochemistry* **33**, 14475 (1994).
- [153] K.K. Surerus, M.C. Kennedy, H. Beinert, and E. Münck: Mössbauer study of the inactive  $\text{Fe}_3\text{S}_4$  and  $\text{Fe}_3\text{Se}_4$  and the active  $\text{Fe}_4\text{Se}_4$  forms of beef-heart aconitase. *Proc. Natl. Acad. Sci. U.S.A.* **86**, 9846 (1989).
- [154] K.K. Surerus, M. Chen, J.W. van der Zwaan, F.M. Rusnak, M. Kolk, E.C. Duin, S.P.J. Albracht, and E. Münck: Further Characterization of the Spin Coupling Observed in Oxidized Hydrogenase from *Chromatium vinosum*. A Moessbauer and Multifrequency EPR Study. *Biochemistry* **33**, 4980 (1994).
- [155] L. Noodleman and D.A. Case: Density functional theory of spin polarization and spin coupling in iron-sulfur clusters. *Adv. Inorg. Chem.* **38**, 423 (1992).
- [156] J.-M. Mouesca, L. Noodleman, D.A. Case, and B. Lamotte: Spin Densities and Spin Coupling in Iron-Sulfur Clusters: A New Analysis of Hyperfine Coupling Constants. *Inorg. Chem.* **34**, 4347 (1995).
- [157] J.A. Pople, M. Head-Gordon, D.J. Fox, K. Raghavachari, and L.A. Curtiss: Gaussian-1 theory: A general procedure for prediction of molecular energies. *J. Chem. Phys.* **90**, 5622 (1989).
- [158] P.E. Blöchl: Electrostatic Decoupling of Periodic Images of Plane-Wave-Expanded densities and Derived Atomic Point Charges. *J. Chem. Phys.* **103**, 7422 (1986).
- [159] R. Hoffmann: *A chemist's view of bonding in extended structures* VCH Publishers, New York (1988).
- [160] A. F. Holleman and E. Wiberg: *Lehrbuch der Anorganischen Chemie* Gruyter (1995).
- [161] H. Petrilli, P.E. Blöchl, P. Blaha, and K. Schwarz: Electric-field-gradient calculations using the projector augmented wave method. *Phys. Rev. B* **57**, 14690–14697 (1998).

- [162] P. Dufek, P. Blaha, and K. Schwarz: Determination of the Nuclear Quadrupole Moment of  $^{57}\text{Fe}$ . *Phys. Rev. Lett.* **75**, 3545 (1995).
- [163] M. C. Durrant: An Atomic Level Model for the Interactions of Molybdenum Nitrogenase with Carbon Monoxide, Acetylene, and Ethylene. *Biochemistry* **43**, 6030 (2004).
- [164] F. Barrière: Modeling of the molybdenum center in the nitrogenase FeMo-cofactor. *Coordin. Chem. Rev.* **236**, 71 (2003).
- [165] D. Sellmann and J. Sutter: Elementary reactions, structure-function relationships, and the potential relevance of low molecular metal-sulfur ligand complexes to biological  $\text{N}_2$  fixation. *J. Biol. Inorg. Chem. (JBIC)* **1**, 597 (1996).
- [166] D. Sellmann and J. Sutter: In Quest of Competitive Catalysts for Nitrogenases and Other Metal Sulfur Enzymes. *Acc. Chem. Res.* **30**, 460 (1997).
- [167] D. Sellmann, J. Utz, N. Blum, and F.H. Heinemann: On the function of nitrogenase FeMo cofactors and competitive catalysts: chemical principles, structural blue-prints, and the relevance of iron sulfur complexes for  $\text{N}_2$  fixation. *Coord. Chem. Rev.* **190-192**, 607 (1999).
- [168] D. Sellmann, A. Fürsattel, and J. Sutter: The nitrogenase catalyzed  $\text{N}_2$  dependent HD formation: a model reaction and its significance for the FeMoco function. *Coord. Chem. Rev.* **200-202**, 545 (2000).
- [169] J.J. Mortensen, L.B. Hansen, B. Hammer, and J.K. Nørskov: Nitrogen Adsorption and Dissociation on Fe(111). *J. Catal.* **182**, 479 (1999).
- [170] A.K. Rappé, C.J. Casewit, K.S. Colwell, W.A. Goddard, and W.M. Skiff: UFF, a full periodic table force field for molecular mechanics and molecular dynamics simulations. *J. Am. Chem. Soc.* **114**, 10024 (1992).
- [171] J.A. Pool, E. Lobkovsky, and P.J. Chirik: Hydrogenation and cleavage of dinitrogen to ammonia with a zirconium complex. *Nature* **427**, 527 (2004).
- [172] D.V. Yandulov and R.R. Schrock: Catalytic Reduction of Dinitrogen to Ammonia at a Single Molybdenum Center. *Science* **301**, 76–78 (2003).
- [173] R.A. Alberty: Thermodynamics of the nitrogenase reactions. *J. Biol. Chem.* **269**, 7099 (1994).
- [174] M.J. Dilworth: Acetylene reduction by nitrogen-fixing preparations from *Clostridium pasteurianum*. *Biochim. Biophys. Acta* **127**, 285 (1966).

- 
- [175] K. Fisher, M. J. Dilworth, C.-H. Kim, and W. E. Newton: *Azotobacter vinelandii* Nitrogenases Containing Altered MoFe Proteins with Substitutions in the FeMo-Cofactor Environment: Effects on the Catalyzed Reduction of Acetylene and Ethylene. *Biochemistry* **39**, 2970 (2000).
- [176] P. M. C. Benton, J. Christiansen, D. R. Dean, and L. C. Seefeldt: Stereospecificity of Acetylene Reduction Catalyzed by Nitrogenase. *J. Am. Chem. Soc.* **123**, 1822 (2001).
- [177] P.A. McLean, A. True, M.J. Nelson, H.-I. Lee, B.M. Hoffman, and W.H. Orme-Johnson: Effects of substrates (methyl isocyanide,  $C_2H_2$ ) and inhibitor (CO) on resting-state wild-type and NifV<sup>-</sup> *Klebsiella pneumoniae* MoFe proteins. *J. Inorg. Biochem.* **93**, 18 (2002).
- [178] D.J. Lowe, K. Fisher, and Thorneley R.N.: *Klebsiella pneumoniae* nitrogenase. Mechanism of acetylene reduction and its inhibition by carbon monoxide. *Biochem. J.* **272**, 621 (1990).
- [179] J.C. Hwang, C.H. Chen, and R.H. Burris: Inhibition of nitrogenase-catalyzed reductions. *Biochim. Biophys. Acta* **292**, 256 (1973).
- [180] L.C. Davis and Y-L. Wang: In Vivo and In Vitro Kinetics of Nitrogenase. *J. Bacteriol.* **141**, 1230 (1980).
- [181] J. Liang and R.H. Burris: Interactions among  $N_2$ ,  $N_2O$ , and  $C_2H_2$  as substrates and inhibitors of nitrogenase from *Azotobacter vinelandii*. *Biochemistry* **27**, 6726 (1988).
- [182] M. Sørli, J. Christiansen, D.R. Dean, and B.J. Hales: Detection of a New Radical and FeMo-Cofactor EPR Signal during Acetylene Reduction by the  $\alpha$ -H195Q Mutant of Nitrogenase. *J. Am. Chem. Soc.* **121**, 9457 (1999).

# Appendix





# A Appendix

## A.1 Total spin of a Slater determinant

The expectation value of the total spin operator  $\mathcal{S}^2$  of a Slater determinant, such as the Kohn-Sham wave function, is used in section 5.5.4 on page 58. With  $|i\rangle$  and  $|j\rangle$  being orthonormalized one-particle states of the Slater determinant,  $\mathcal{S}^2$  can be calculated as follows:

$$\langle \mathcal{S}^2 \rangle = \sum_{ij} \langle \vec{\mathcal{S}}_i \vec{\mathcal{S}}_j \rangle = \sum_i \langle \vec{\mathcal{S}}_i \vec{\mathcal{S}}_i \rangle + \sum_{i \neq j} \langle \vec{\mathcal{S}}_i \vec{\mathcal{S}}_j \rangle \quad (\text{A.1})$$

$$= \sum_i \langle i | \mathcal{S}^2 | i \rangle + \sum_{i \neq j} \left( \langle i | \vec{\mathcal{S}} | i \rangle \langle j | \vec{\mathcal{S}} | j \rangle - \langle i | \vec{\mathcal{S}} | j \rangle \langle j | \vec{\mathcal{S}} | i \rangle \right) \quad (\text{A.2})$$

$$= \sum_i \langle i | \mathcal{S}^2 | i \rangle + \sum_{ij} \left( \langle i | \vec{\mathcal{S}} | i \rangle \langle j | \vec{\mathcal{S}} | j \rangle - \langle i | \vec{\mathcal{S}} | j \rangle \langle j | \vec{\mathcal{S}} | i \rangle \right) \quad (\text{A.3})$$

$$= \sum_{ij} \langle i | \vec{\mathcal{S}} | i \rangle \langle j | \vec{\mathcal{S}} | j \rangle + \sum_i \langle i | \mathcal{S}^2 | i \rangle - \sum_{ij} \langle i | \vec{\mathcal{S}} | j \rangle \langle j | \vec{\mathcal{S}} | i \rangle \quad (\text{A.4})$$

$$= \sum_{ij} \langle i | \vec{\mathcal{S}} | i \rangle \langle j | \vec{\mathcal{S}} | j \rangle + \sum_i \langle i | \vec{\mathcal{S}} \left( 1 - \sum_j |j\rangle \langle j| \right) \vec{\mathcal{S}} | i \rangle \quad (\text{A.5})$$

$$= \left( \sum_i \langle i | \vec{\mathcal{S}} | i \rangle \right)^2 + \sum_i \langle i | \vec{\mathcal{S}} \left( 1 - \sum_j |j\rangle \langle j| \right) \vec{\mathcal{S}} | i \rangle \quad (\text{A.6})$$

The first term is the square of the integrated spin density

$$\left( \sum_i \langle i | \vec{\mathcal{S}} | i \rangle \right)^2 = \left( \frac{\hbar}{2} \int \vec{n}_s(\vec{r}) d^3r \right)^2 \quad (\text{A.7})$$

The second term in (A.6) is an exchange-like term that can be evaluated as

follows. From (A.4) one obtains

$$\sum_i \langle i | \vec{S} \left( 1 - \sum_j |j\rangle \langle j| \right) \vec{S} | i \rangle = \sum_i \langle i | \mathcal{S}^2 | i \rangle - \sum_{ij} \langle i | \vec{S} | j \rangle \langle j | \vec{S} | i \rangle \quad (\text{A.8})$$

$$\sum_i \langle i | \mathcal{S}^2 | i \rangle = \sum_i \frac{3\hbar^2}{4} = \frac{3N\hbar^2}{4} \quad (\text{A.9})$$

$$\sum_{ij} \langle i | \vec{S} | j \rangle \langle j | \vec{S} | i \rangle = \sum_{ij} \left( \langle i | \mathcal{S}_x | j \rangle \langle j | \mathcal{S}_x | i \rangle + \langle i | \mathcal{S}_y | j \rangle \langle j | \mathcal{S}_y | i \rangle + \langle i | \mathcal{S}_z | j \rangle \langle j | \mathcal{S}_z | i \rangle \right). \quad (\text{A.10})$$

Inserting the spinor components and the Pauli matrix (5.67) one obtains

$$\langle i | \mathcal{S}_z | j \rangle = \frac{\hbar}{2} (\langle i_\uparrow | j_\uparrow \rangle - \langle i_\downarrow | j_\downarrow \rangle) \quad (\text{A.11})$$

and the  $\mathcal{S}_z$  term becomes

$$\begin{aligned} \langle i | \mathcal{S}_z | j \rangle \langle j | \mathcal{S}_z | i \rangle &= \frac{\hbar^2}{4} (\langle i_\uparrow | j_\uparrow \rangle - \langle i_\downarrow | j_\downarrow \rangle) (\langle j_\uparrow | i_\uparrow \rangle - \langle j_\downarrow | i_\downarrow \rangle) = \\ &= \frac{\hbar^2}{4} |\langle i_\uparrow | j_\uparrow \rangle - \langle i_\downarrow | j_\downarrow \rangle|^2. \end{aligned} \quad (\text{A.12})$$

Using

$$\mathcal{S}_+ = \mathcal{S}_x + i\mathcal{S}_y, \quad \mathcal{S}_- = \mathcal{S}_x - i\mathcal{S}_y \quad (\text{A.13})$$

$$\mathcal{S}_x = \frac{1}{2}(\mathcal{S}_+ + \mathcal{S}_-), \quad \mathcal{S}_y = \frac{-i}{2}(\mathcal{S}_+ - \mathcal{S}_-), \quad (\text{A.14})$$

$$\langle i | \mathcal{S}_+ | j \rangle = \hbar \langle i_\uparrow | j_\downarrow \rangle \quad (\text{A.15})$$

$$\langle i | \mathcal{S}_- | j \rangle = \hbar \langle i_\downarrow | j_\uparrow \rangle \quad (\text{A.16})$$

$$(\text{A.17})$$

one obtains for the  $\mathcal{S}_x$  and  $\mathcal{S}_y$  terms

$$\langle i | \mathcal{S}_x | j \rangle \langle j | \mathcal{S}_x | i \rangle = \frac{1}{4} \left( \langle i | \mathcal{S}_+ | j \rangle + \langle i | \mathcal{S}_- | j \rangle \right) \left( \langle j | \mathcal{S}_+ | i \rangle + \langle j | \mathcal{S}_- | i \rangle \right) \quad (\text{A.18})$$

$$\langle i | \mathcal{S}_y | j \rangle \langle j | \mathcal{S}_y | i \rangle = -\frac{1}{4} \left( \langle i | \mathcal{S}_+ | j \rangle - \langle i | \mathcal{S}_- | j \rangle \right) \left( \langle j | \mathcal{S}_+ | i \rangle - \langle j | \mathcal{S}_- | i \rangle \right) \quad (\text{A.19})$$

$$\begin{aligned} \langle i | \mathcal{S}_x | j \rangle \langle j | \mathcal{S}_x | i \rangle + \langle i | \mathcal{S}_y | j \rangle \langle j | \mathcal{S}_y | i \rangle &= \frac{1}{2} \langle i | \mathcal{S}_+ | j \rangle \langle j | \mathcal{S}_- | i \rangle + \frac{1}{2} \langle i | \mathcal{S}_- | j \rangle \langle j | \mathcal{S}_+ | i \rangle = \\ &= \frac{\hbar^2}{2} \left( \langle i_\uparrow | j_\downarrow \rangle \langle j_\downarrow | i_\uparrow \rangle + \langle i_\downarrow | j_\uparrow \rangle \langle j_\uparrow | i_\downarrow \rangle \right) = \frac{\hbar^2}{2} \left( |\langle i_\uparrow | j_\downarrow \rangle|^2 + |\langle i_\downarrow | j_\uparrow \rangle|^2 \right). \end{aligned} \quad (\text{A.20})$$

$$\sum_{ij} \left( |\langle i_\uparrow | j_\downarrow \rangle|^2 + |\langle i_\downarrow | j_\uparrow \rangle|^2 \right) = 2 \sum_{ij} |\langle i_\uparrow | j_\downarrow \rangle|^2 \quad (\text{A.21})$$

This results in a second term of equation (A.6) as

$$\sum_i \langle i | \vec{S} \left( 1 - \sum_j |j\rangle \langle j| \right) \vec{S} | i \rangle = \frac{\hbar^2}{4} \sum_{ij} \left( 3\delta_{ij} - |\langle i_\uparrow | j_\uparrow \rangle - \langle i_\downarrow | j_\downarrow \rangle|^2 - 4|\langle i_\uparrow | j_\downarrow \rangle|^2 \right) \quad (\text{A.22})$$

Inserting (A.7) and (A.22) in (A.6) it is possible to calculate the expectation value of the total spin as two-particle operator.

$$\langle \mathcal{S}^2 \rangle = \underbrace{\left( \frac{\hbar}{2} \int \vec{n}_s(\vec{r}) d^3r \right)^2}_{\text{classical}} + \underbrace{\frac{\hbar^2}{4} \sum_{ij} \left( 3\delta_{ij} - |\langle i_\uparrow | j_\uparrow \rangle - \langle i_\downarrow | j_\downarrow \rangle|^2 - 4|\langle i_\uparrow | j_\downarrow \rangle|^2 \right)}_{\text{quantum-mechanical}} \quad (\text{A.23})$$

## A.2 Total energies of intermediates

state	E[H]	#H	#NH <sub>3</sub>	E <sub>rel</sub>
<i>M</i>	-398.35889	-2	0	0
<i>MH</i>	-398.92886	-1	0	0
<i>MH2</i>	-399.49406	0	0	13
<i>A0</i>	-399.50550	0	0	-18
<i>A1</i>	-399.50333	0	0	-12
<i>A2</i>	-399.48547	0	0	35
<i>A3</i>	-399.47217	0	0	70
<i>B0</i>	-400.07180	1	0	-8
<i>B1</i>	-400.05765	1	0	29
<i>B1'</i>	-400.05476	1	0	37
<i>B2</i>	-400.05038	1	0	48
<i>B3</i>	-400.04880	1	0	53
<i>B4</i>	-400.04687	1	0	58
<i>B5</i>	-400.04573	1	0	61
<i>B6</i>	-400.03945	1	0	77
<i>B7</i>	-400.01668	1	0	137
<i>C0</i>	-400.65651	2	0	-47
<i>C1</i>	-400.64889	2	0	-27
<i>C2</i>	-400.64340	2	0	-12
<i>C3</i>	-400.63860	2	0	0
<i>C4</i>	-400.62061	2	0	48
<i>C01</i>	-400.62613	2	0	33
<i>D0</i>	-401.24422	3	0	-93
<i>D1</i>	-401.23728	3	0	-75
<i>D2</i>	-401.23685	3	0	-74
<i>E0</i>	-390.02906	4	1	-246
<i>E1</i>	-401.84055	4	0	-162
<i>E2</i>	-401.79484	4	0	-42
<i>E3</i>	-401.78779	4	0	-24
<i>F0</i>	-390.65556	5	1	-395
<i>F1</i>	-390.64083	5	1	-356
<i>F2</i>	-390.63626	5	1	-344
<i>F3</i>	-390.62317	5	1	-310
<i>F4</i>	-390.61492	5	1	-288
<i>F5</i>	-402.36783	5	0	-50

Total energies of intermediates of the reaction mechanism. E[H] is the total energy in hartrees, #H the number of transferred hydrogen atoms with respect to the resting state, #NH<sub>3</sub> the number of ammonia molecules separated from the cofactor complex and E<sub>rel</sub> the relative energy in kJ/mol of the corresponding intermediate using a  $\mu_H$  to adjust *M* and *MH* to the same energy as described in section 10.3 on page 128.

## B CV of Johannes Kästner

### Personal Data

Day of Birth: 1978-01-30  
Place of Birth: Vienna, Austria  
Birth Name: Johannes Schimpl  
Sex: M  
Marital status: Married (2004-04-05)  
to Mag. Andrea Kästner  
Parents' Names: Dipl. Ing. Karl and  
Mag. Monika Schimpl  
Nationality: Austrian  
Language competencies: German (native language),  
English (excellent),  
French (medium)



### Education

2001 - present Scientific coworker with Peter E. Blöchl at the Institute of Theoretical Physics, working on my Ph.D.  
2001-07-02 Master's degree with honors in chemistry  
2000 - 2001 Diploma thesis in chemistry, K.H. Schwarz, advisor, Vienna University of Technology, Institute of Physical and Theoretical Chemistry (now Institute of Materials Chemistry). Title: "Linkage Isomers of Transition Metal Nitrosyl Complexes studied by ab-initio Molecular Dynamics"  
1996 - 2001 Study of chemistry, Vienna University of Technology, 1997 - 2001 education in physics, 1999-05-25 completion of undergraduate study in chemistry with honors  
1988 - 1996 High school: BRG Oberschützen, Austria, thesis in chemistry at the final exams (passed 1996-06-18)

**Awards and Honors**

2002	Prize of the city of Vienna for excellent diploma thesis' (730 €)
1996	Award for the thesis at the school-leaving exam from the Austrian Chemical Society (GÖCH)
1996	Gold medal and Sorokin special award, International Chemistry Olympiad (IChO), Moscow
1995	Gold medal, International Chemistry Olympiad (IChO), Beijing
1995 and 1996	Gold medal, National Chemistry Olympiad (Austria)

**Appointments**

2001 - present	Research Assistant with Peter E. Blöchl Institute of Theoretical Physics, Clausthal University of Technology (Germany)
1999 - 2001	Instructor of problem sessions, Institute of Experimental Physics, Vienna University of Technology
1999	Process optimization in synthetic chemistry, internship (4 weeks), "ÖCW der Degussa Austria", Weißenstein (Austria)
1998	Hard metal works, internship (4 weeks), "Plansee-Tizit", Reutte (Austria)

**Publications**

- "Nitrogen fixation at the FeMo-cofactor of nitrogenase: A comprehensive DFT investigation of reaction pathways and their energetics.", Johannes Kästner and Peter E. Blöchl, submitted
- "Electronic structure methods: Augmented Waves, Pseudopotentials and the Projector Augmented Wave method", Peter E. Blöchl, Clemens J. Först and Johannes Kästner, Book chapter in "Handbook of Materials Modeling" Vol. 1, R. Catlow, H. Shercliff and S. Yip Eds. (Kluwer Academic Publishers), scheduled for 2004

- “Nitrogen Binding to the FeMo-Cofactor of Nitrogenase”, Johannes Schimpl, Helena M. Petrilli and Peter E. Blöchl, *J. Am. Chem. Soc.* **125**, 15772 (2003)
- “Projector augmented wave method: ab initio molecular dynamics with full wave functions”, Peter E. Blöchl, Clemens J. Först and Johannes Schimpl, *Bull. Mater. Sci.* **26**, p33 (2003)

### Teaching

#### Clausthal:

2003	Problem sessions theoretical physics 2 - electrodynamics
2003	Physics laboratory
2002/03	Problem sessions theoretical physics 1 - classical mechanics
2001/02	Problem sessions theoretical physics 3 - quantum mechanics

#### Vienna:

2000/01	Problem sessions “quantum mechanics - introduction”
2000/01	Problem sessions “thermodynamics”
2000	Problem sessions 2 for physics students
1999/00	Problem sessions 1 for physics students

### Conferences

- Gordon Research Conference on “Electron Distribution & Chemical Bonding”: Mount Holyoke College, South Hadley, MA, USA, July 4-9, 2004. Invited talk: Johannes Kästner, P.E. Blöchl: *Unraveling biological nitrogen fixation: Insights from first-principles calculations.*
- Gordon Research Conference on “Nitrogen Fixation”: Colby-Sawyer College, New London, NH, USA, June 20-25 2004. Invited talk: Johannes Kästner, P.E. Blöchl: *Active role of the FeMoco cage: a comprehensive model of the mechanism from first-principles calculations.*
- “19. Workshop on Novel Materials and Superconductors: Planneralp”, February 22. - 28., 2004. Poster presentation: Johannes Schimpl and Peter E. Blöchl: *Unraveling biological nitrogen fixation: Insights from first-principles calculations*

- *Reaction mechanism of the enzyme nitrogenase*: Invited talk at the Institute for Physical and Theoretical Chemistry, Bonn University, December 11 2003.
- “Workshop on Application of Density-Functional Theory in Condensed Matter Physics, Surface Physics, Chemistry, Engineering and Biology”: Fritz-Haber-Institut of Max-Planck-Gesellschaft, Berlin, Germany, July 21-30 2003. Poster presentation: Johannes Schimpl, P.E. Blöchl: *Nitrogen Adsorption at the Cofactor of Nitrogenase*
- Symposium “Computational Modelling of Catalysis”: Max Planck Institut für Kohlenforschung, Mülheim an der Ruhr, Germany, July 16-18 2003. Poster presentation: Johannes Schimpl, P.E. Blöchl: *Nitrogen Adsorption at the Cofactor of Nitrogenase*
- HLRN-Workshop “Quantum chemistry/Density functional theory”: Berlin, July 8 2003. Talk: J. Schimpl: *PAW-Rechnungen auf dem HLRN-System und PC-Clustern*
- “18. Workshop on Novel Materials and Superconductors”, Planneralp, February 23 - march 1 2003
- “17. Workshop on Novel Materials and Superconductors”, Planneralp, February 24 - march 2 2002. Contributed talk: Johannes Schimpl, P.E. Blöchl: *Non-Collinear Magnetism of Iron-Sulfur Proteins*
- 80th Bunsen-Colloquium Stuttgart 2001; main topic: computer simulations in physical chemistry: large molecules, liquids, solid state, may 24-26 2001
- “16. Workshop on Novel Materials and Superconductors”: Planneralp, 2001. Poster presentation: Johannes Schimpl and Karlheinz Schwarz, *Molecular dynamics study of nitroprussides  $\text{Na}_2[\text{Fe}(\text{CN})_5\text{NO}]$  used in holography*
- “Applied DFT2001”: International Conference on Applied Density Functional Theory, 14.-17.1.2001 Vienna



## C Papers

- “Nitrogen Binding to the FeMo-Cofactor of Nitrogenase”, Johannes Schimpl, Helena M. Petrilli and Peter E. Blöchl, *J. Am. Chem. Soc.* **125**, 15772 (2003)
- “Nitrogen fixation at the FeMo-cofactor of nitrogenase: A comprehensive DFT investigation of reaction pathways and their energetics.”, Johannes Kästner and Peter E. Blöchl, submitted